Quantification of LV function and mass by cardiovascular magnetic resonance: multi-center variability and consensus contours

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

Citation

Published Version

Citable link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:21462232

Terms of Use
This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA
Quantification of LV function and mass by cardiovascular magnetic resonance: multi-center variability and consensus contours

Avan Suinesiaputra¹, David A. Bluemke², Brett R. Cowan¹, Matthias G. Friedrich³, Christopher M. Kramer⁴, Raymond Kwong⁵, Sven Plein⁶, Jeanette Schulz-Menger⁸, Jos J. M. Westenberg⁷, Alistair A. Young*¹ and Eike Nagel⁹

Abstract

Background: High reproducibility of LV mass and volume measurement from cine cardiovascular magnetic resonance (CMR) has been shown within single centers. However, the extent to which contours may vary from center to center, due to different training protocols, is unknown. We aimed to quantify sources of variation between many centers, and provide a multi-center consensus ground truth dataset for benchmarking automated processing tools and facilitating training for new readers in CMR analysis.

Methods: Seven independent expert readers, representing seven experienced CMR core laboratories, analyzed fifteen cine CMR data sets in accordance with their standard operating protocols and SCMR guidelines. Consensus contours were generated for each image according to a statistical optimization scheme that maximized contour placement agreement between readers.

Results: Reader-consensus agreement was better than inter-reader agreement (end-diastolic volume 14.7 ml vs 15.2–28.4 ml; end-systolic volume 13.2 ml vs 14.0–21.5 ml; LV mass 17.5 g vs 20.2–34.5 g; ejection fraction 4.2 % vs 4.6–7.5 %). Compared with consensus contours, readers were very consistent (small variability across cases within each reader), but bias varied between readers due to differences in contouring protocols at each center. Although larger contour differences were found at the apex and base, the main effect on volume was due to small but consistent differences in the position of the contours in all regions of the LV.

Conclusions: A multi-center consensus dataset was established for the purposes of benchmarking and training. Achieving consensus on contour drawing protocol between centers before analysis, or bias correction after analysis, is required when collating multi-center results.

Keywords: Left ventricular function, Benchmarks, Reproducibility, Multi-center, Left ventricular mass, Image analysis

Background

Left ventricular (LV) mass and volumes are essential for the management of patients with cardiovascular disease. In particular, LV mass (LVM) is an independent predictor of cardiovascular events [1], and end-diastolic volume (EDV) and end-systolic volume (ESV) are associated with adverse remodeling [2]. Cardiovascular magnetic resonance (CMR) is currently the most accurate and reproducible method for quantifying LV mass and volumes [3]. CMR is non-invasive, does not require geometrical assumptions, is non-ionising and has high signal-to-noise ratio. Because of these advantages, CMR is becoming widely used for the measurement of ventricular volumes, function and mass in many clinical centers, as well as in large research studies including the Multi-Ethnic Study of Atherosclerosis (MESA) [4] and the UK Biobank [5].

LV mass and volume quantification requires accurate delineation of the blood pool and myocardium. Although the contrast between flowing blood and the myocardium

* Correspondence: a.young@auckland.ac.nz
¹Department of Anatomy with Radiology, University of Auckland, Auckland, New Zealand
Full list of author information is available at the end of the article

© 2015 Suinesiaputra et al. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.
in steady-state free precession (SSFP) images is typically excellent, the precise placement of the contours is reader dependent [6]. High reproducibility of LV mass and volume measurement based on cine CMR has been shown within single centers [7, 8], but differences in training and standard operating procedures may occur between centers [6]. A multi-center consensus ground truth dataset would be valuable for evaluating sources of variability, training new readers, establishing standard protocols for multi-centre studies and validating new computer algorithms for automated contouring. Such a dataset is difficult to establish, due to the time-consuming nature of manual contouring. Although there are several recourses that offer large breadth [5, 9, 10], no resource offers depth of expert analysis from multiple centers. Greater depth of readers is valuable in evaluating sources of variation and an unbiased consensus. The aim of this study was to develop a consensus ground truth LV contour dataset for SSFP cine images, derived from expert readers representing seven independent CMR centers from many countries around the world, in accordance with the SCMR post-processing guidelines [11].

Methods
Participants
Cine CMR images from 15 subjects (five healthy volunteers, six patients with myocardial infarction, two patients with heart failure, and two patients with LV hypertrophy) were included in this study. CMR images were acquired with contiguous short axis slices and 2–3 long axis slices in accordance with SCMR guidelines using three different scanners (4 GE, 5 Siemens and 6 Philips). Spatial resolution varied with FOV, ranging from 92 × 72 to 280 × 280 mm² (see Table 1). Temporal resolution was typically 20–30 frames, except for one case with 60 cardiac frames. Slice thickness was either 8 or 10 mm. Short-axes view series covering the LV from apex to base were defined in 10–15 slices. Anonymized images were contributed to the Cardiac Atlas Project database with the approval of local institutional review boards. Written informed consent was obtained from all participants.

Analysis
Seven expert readers, representing seven CMR core laboratories from six countries around the world (two USA, Canada, two UK, Germany and Netherlands) analyzed all cases independently. The contours were reviewed by the core laboratory principal investigator and represented the standard practice of each core laboratory. Each center was able to use their usual software, with the condition that all contours were placed manually, or manually corrected if initial contours were found automatically. Contouring was performed in accordance with the SCMR guidelines on standardized image interpretation and post-processing of CMR images [11]. Trabeculae and papillary muscles were included in the blood pool and excluded from the LV mass. No attempt was made to train the readers, influence their analysis or achieve consistent results between readers.

Epicardial and endocardial contours were drawn on the short-axis slices at end-diastole (ED) and endocardial contours drawn at end-systole (ES). The ED and ES frames

<table>
<thead>
<tr>
<th>Case</th>
<th>Pathology</th>
<th>Gender</th>
<th>Age</th>
<th>Slices</th>
<th>Frames</th>
<th>SLT</th>
<th>SLD</th>
<th>FOV</th>
<th>Pixel size</th>
<th>Vendor</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Healthy</td>
<td>M</td>
<td>70</td>
<td>10</td>
<td>20</td>
<td>10</td>
<td>10</td>
<td>205 × 205</td>
<td>1.25 × 1.25</td>
<td>GE</td>
</tr>
<tr>
<td>2</td>
<td>Infarct</td>
<td>M</td>
<td>50</td>
<td>13</td>
<td>25</td>
<td>6</td>
<td>10</td>
<td>92 × 75</td>
<td>2.08 × 2.08</td>
<td>GE</td>
</tr>
<tr>
<td>3</td>
<td>Heart Failure</td>
<td>F</td>
<td>77</td>
<td>15</td>
<td>20</td>
<td>8</td>
<td>10</td>
<td>211 × 211</td>
<td>1.21 × 1.21</td>
<td>GE</td>
</tr>
<tr>
<td>4</td>
<td>Infarct</td>
<td>M</td>
<td>70</td>
<td>12</td>
<td>27</td>
<td>6</td>
<td>10</td>
<td>182 × 137</td>
<td>1.41 × 1.41</td>
<td>SIEMENS</td>
</tr>
<tr>
<td>5</td>
<td>Heart Failure</td>
<td>F</td>
<td>63</td>
<td>14</td>
<td>20</td>
<td>8</td>
<td>8</td>
<td>187 × 187</td>
<td>1.37 × 1.37</td>
<td>GE</td>
</tr>
<tr>
<td>6</td>
<td>Hypertrophy</td>
<td>M</td>
<td>61</td>
<td>12</td>
<td>20</td>
<td>10</td>
<td>10</td>
<td>187 × 187</td>
<td>1.37 × 1.37</td>
<td>GE</td>
</tr>
<tr>
<td>7</td>
<td>Infarct</td>
<td>M</td>
<td>59</td>
<td>11</td>
<td>25</td>
<td>6</td>
<td>10</td>
<td>102 × 102</td>
<td>1.88 × 1.88</td>
<td>SIEMENS</td>
</tr>
<tr>
<td>8</td>
<td>Infarct</td>
<td>F</td>
<td>42</td>
<td>11</td>
<td>24</td>
<td>6</td>
<td>10</td>
<td>182 × 142</td>
<td>1.41 × 1.41</td>
<td>GE</td>
</tr>
<tr>
<td>9</td>
<td>Infarct</td>
<td>M</td>
<td>52</td>
<td>11</td>
<td>22</td>
<td>6</td>
<td>10</td>
<td>182 × 137</td>
<td>1.41 × 1.41</td>
<td>GE</td>
</tr>
<tr>
<td>10</td>
<td>Healthy</td>
<td>M</td>
<td>67</td>
<td>15</td>
<td>30</td>
<td>8</td>
<td>8</td>
<td>208 × 208</td>
<td>1.23 × 1.23</td>
<td>GE</td>
</tr>
<tr>
<td>11</td>
<td>Healthy</td>
<td>F</td>
<td>66</td>
<td>11</td>
<td>60</td>
<td>10</td>
<td>10</td>
<td>177 × 177</td>
<td>1.45 × 1.45</td>
<td>SIEMENS</td>
</tr>
<tr>
<td>12</td>
<td>Healthy</td>
<td>M</td>
<td>50</td>
<td>13</td>
<td>30</td>
<td>10</td>
<td>10</td>
<td>156 × 156</td>
<td>1.44 × 1.44</td>
<td>SIEMENS</td>
</tr>
<tr>
<td>13</td>
<td>Infarct</td>
<td>M</td>
<td>70</td>
<td>15</td>
<td>30</td>
<td>8</td>
<td>8</td>
<td>280 × 280</td>
<td>1.14 × 1.14</td>
<td>SIEMENS</td>
</tr>
<tr>
<td>14</td>
<td>Healthy</td>
<td>M</td>
<td>57</td>
<td>15</td>
<td>30</td>
<td>8</td>
<td>8</td>
<td>277 × 277</td>
<td>1.21 × 1.21</td>
<td>PHILIPS</td>
</tr>
<tr>
<td>15</td>
<td>Hypertrophy</td>
<td>M</td>
<td>71</td>
<td>15</td>
<td>30</td>
<td>8</td>
<td>8</td>
<td>241 × 241</td>
<td>1.19 × 1.19</td>
<td>PHILIPS</td>
</tr>
</tbody>
</table>
were pre-determined for all readers by the coordinating center, based on smallest area in a mid-ventricular slice. All readers were asked to contour slices covering the whole ventricle from apex to base, but there was no restriction on which slices to include or exclude.

Readers used a range of software packages. Two readers used Osirix (Pixmeo, Geneva, Switzerland), five readers used QMass (Medis, Leiden, the Netherlands), and two readers used CMR42 (Circle Cardiovascular Imaging Inc., Calgary, Canada). Two readers used two different software packages. Contours were imported from these software packages and pre-processed using Matlab R2010a (Mathworks, Natick, MA, USA). Consensus contours were only generated if most of the readers (i.e. four or more) contoured the slice; otherwise, no consensus contours were produced.

Consensus contour estimation
Consensus contours were estimated using the Simultaneous Truth and Performance Level Estimation (STAPLE) method [12]. This method calculated unique contours in each slice that maximized the conditional probability of the consensus given the readers’ contours. Briefly, contours from each reader were first converted to binary images (1 for pixels within the contour, 0 for pixels outside the contour). Since the contour resolution was higher than the original images, the binary images were calculated at a resolution 4× higher than the original image. Given an estimate of the consensus contours, the sensitivity of each reader was calculated as the proportion of pixels inside the consensus contours that were also inside the reader contours. Similarly, the reader specificity was calculated as the proportion of pixels outside the consensus contours, which were also outside the reader contours. The STAPLE method uses Expectation-Maximization [13] to calculate the optimal consensus contour, as well as the reader sensitivity and specificity. There are two steps that are performed iteratively until convergence. The first step (Expectation) estimates the consensus probability given the reader contours and current estimates of sensitivity and specificity. The second step (Maximization) updates the sensitivity and specificity of each reader based on this consensus probability. The result is not the same as simple averaging or pixel voting, since specificity and sensitivity behave as weights during the optimization process, which are not assumed to be equal across all readers. Instead, the voting solution is used as the initial estimate to start the iteration. The STAPLE method has been successfully applied in several medical imaging applications, and was recently used to estimate consensus ground truth contours from automated CMR analysis methods [9].

Cavity volumes and myocardial mass
LV cavity volumes at ED (EDV) and ES (ESV) were computed by slice summation. Two-dimensional cavity areas were multiplied by the inter slice distance to compute a slice volume. The myocardial mass (LVM) was defined at ED by subtracting EDV from epicardial volume and multiplying by 1.05 g/ml.

Functional assessment
Root mean squared errors (RMSE) in volumes and mass were computed to measure the agreement between each reader and all the other readers. This was defined as

$$E_i(F) = \sqrt{\frac{\sum_{j=1}^{R} \sum_{k=1}^{N}(F_j(k) - F_i(k))^2}{N(R-1)}}$$

where $E_i$ indicates the RMSE for reader $i$, $j$ indicates all other readers, $F$ indicates either EDV, ESV, LVM or ejection fraction (EF), $R$ is the number of readers, $k$ indicates the cases, and $N$ is the number of cases. A similar RMSE was applied also to the consensus, denoted by $E_C(F)$, to measure the agreement between the consensus and all readers:

$$E_C(F) = \sqrt{\frac{\sum_{j=1}^{R} \sum_{k=1}^{N}(F_C(k) - F_i(k))^2}{NR}}$$

Smaller values of $E_C(F)$ compared to $E_i(F)$ for all $i = 1, 2, \ldots, R$ were taken to indicate functionally acceptable consensus contours.

Visual assessment
An independent reader with over 10 years experience in CMR, who was not affiliated with any of the participating core laboratories, visually assessed the consensus contours by scoring as either acceptable or unacceptable according to whether the contour was clinically plausible.

Statistics
Bland-Altman analysis was performed to evaluate variation in volumes and mass across all readers relative to the consensus. The limits of agreement were defined at 95% of confidence interval from the bias. Individual reader bias was quantified by the mean of the differences from the consensus, and reader precision was quantified by the standard deviation of the differences. Volumes and mass estimated from the consensus contours were calculated with the standard error estimated between the readers and the consensus. Statistical analysis was performed using the open source R statistics package (The R Foundation of Statistical Computing Platform, ver. 3.1.1).
Results

Visual assessment
Of all 15 cases, no unacceptable consensus contours were found by the independent reader. Figure 1 shows a representative case of consensus contours estimated from this study. Although large disagreements between readers could be found in some of the apex, base and outflow tract slices, the consensus contours generated from these difficult slices (shown by Fig. 2) were visually acceptable.

Functional assessment
Table 2 shows the estimated consensus LV function per case in terms of EDV, ESV, LVM and EF (ejection fraction). Standard errors between the readers and the consensus were used to indicate confidence intervals for the estimated values.

The RMSE values from the consensus ($E_C$) were always the smallest compared to any reader RMSE values ($E_i$). For EDV, $E_C$ was 14.7 ml while $E_i$ was from 15.2 to 28.4 ml. For ESV, $E_C$ was 13.2 ml while $E_i$ was from 14 to 21.5 ml. For LVM, $E_C$ was 17.5 g, while $E_i$ was from 20.2 to 34.5 g. For EF, $E_C$ was 4.2 % while $E_i$ was from 4.6 to 7.5 %.

Bias and precision
Figure 3 shows reader bias and precision using the estimated consensus values (Table 2) as the reference. All readers showed good precision, with low standard deviations of the differences. The precision in EDV ranged from 5.0 to 11.6 ml (average 8.9 ml); precision in ESV ranged from 7.0 to 11.8 ml (average 9.5 ml); precision in LVM ranged from 10.0 to 12.9 g (average 10.9 g); precision in EF ranged from 1.5 to 5.6 % (average 3.6 %). However, differences in analysis protocols between readers were evident, with some readers exhibiting smaller endocardial and larger epicardial contours, whilst others showed larger endocardial and smaller epicardial contours. The bias in EDV ranged from −36.6 to 40.5 ml (average 0.9 ml); bias in ESV ranged from −32.9 to 41.2 ml (average 0.8 ml); bias in LVM ranged from −44.5 to 59.6 g (average 0.7 g); bias in EF ranged from −11.7 to 12.8 % (average 0.0 %).

Individual reader reports were generated automatically. Figure 4 shows an example report demonstrating similarity of the reader contour with the consensus, with the disparity between expert readers (dark red bands).

Discussion
Readers were consistent within themselves, which was indicated by the relatively small standard deviations (precision) from the consensus within each reader (Fig. 3). However, different readers could be larger or smaller than the consensus (bias). This was due primarily to differing practices at each core lab, where contouring
was consistently smaller or larger compared with the other readers. We also examined other sources of possible bias, in particular in the outflow and apical slices (Fig. 2), but the bias contributions from these slices to the final mass and volume estimate were insignificant.

For clinical studies collating results across core labs, steps should be taken to reduce bias between core labs, either prior to the analysis (by training) or after analysis (by bias correction). This study provides a standard set of cases that could be used as a training set. Alternatively, automated post-hoc bias correction methods [14] can also be applied.

Consensus contour quality
The consensus LV function estimates had the better agreement with all readers than any individual reader. All the RMSE values for EDV, ESV, mass and EF were smallest for the consensus compared with any reader. This indicated that the consensus had better agreement with the readers than any individual reader.

As demonstrated by Fig. 1, the consensus contours were visually acceptable in all cases. Greatest disagreements between readers appeared in the areas where tissue contrast ratios were low, such as in the apical slices and in the outflow tract. Even for these slices, the estimated consensus contours were visually acceptable. Note

<table>
<thead>
<tr>
<th>Case</th>
<th>EDV (ml)</th>
<th>ESV (ml)</th>
<th>LVM (g)</th>
<th>EF (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>104.3 ± 1.8</td>
<td>48.2 ± 2.3</td>
<td>70.4 ± 4.3</td>
<td>53.8 ± 1.5</td>
</tr>
<tr>
<td>2</td>
<td>284.9 ± 6.1</td>
<td>186.7 ± 6.7</td>
<td>154.5 ± 10.0</td>
<td>34.5 ± 1.1</td>
</tr>
<tr>
<td>3</td>
<td>292.8 ± 4.9</td>
<td>253.5 ± 6.4</td>
<td>134.3 ± 7.4</td>
<td>13.4 ± 1.6</td>
</tr>
<tr>
<td>4</td>
<td>190.5 ± 5.0</td>
<td>112.0 ± 4.5</td>
<td>113.7 ± 6.7</td>
<td>41.2 ± 1.2</td>
</tr>
<tr>
<td>5</td>
<td>369.0 ± 9.3</td>
<td>267.5 ± 8.2</td>
<td>171.4 ± 11.8</td>
<td>27.5 ± 0.7</td>
</tr>
<tr>
<td>6</td>
<td>150.6 ± 4.7</td>
<td>47.7 ± 6.4</td>
<td>174.9 ± 5.8</td>
<td>68.3 ± 3.7</td>
</tr>
<tr>
<td>7</td>
<td>190.9 ± 4.8</td>
<td>105.2 ± 4.7</td>
<td>113.5 ± 5.0</td>
<td>44.9 ± 1.2</td>
</tr>
<tr>
<td>8</td>
<td>201.3 ± 4.6</td>
<td>141.6 ± 6.9</td>
<td>122.0 ± 5.7</td>
<td>29.7 ± 2.8</td>
</tr>
<tr>
<td>9</td>
<td>265.1 ± 7.6</td>
<td>160.7 ± 5.4</td>
<td>130.2 ± 5.1</td>
<td>39.4 ± 0.7</td>
</tr>
<tr>
<td>10</td>
<td>157.9 ± 5.6</td>
<td>65.3 ± 4.8</td>
<td>132.7 ± 7.8</td>
<td>58.7 ± 2.0</td>
</tr>
<tr>
<td>11</td>
<td>158.9 ± 4.2</td>
<td>69.1 ± 3.1</td>
<td>99.8 ± 5.2</td>
<td>56.5 ± 1.4</td>
</tr>
<tr>
<td>12</td>
<td>222.2 ± 5.7</td>
<td>88.3 ± 2.4</td>
<td>121.5 ± 7.3</td>
<td>60.3 ± 1.6</td>
</tr>
<tr>
<td>13</td>
<td>216.5 ± 5.3</td>
<td>106.2 ± 5.2</td>
<td>129.1 ± 9.3</td>
<td>50.9 ± 1.5</td>
</tr>
<tr>
<td>14</td>
<td>169.7 ± 4.7</td>
<td>74.3 ± 4.7</td>
<td>116.7 ± 6.1</td>
<td>56.2 ± 2.3</td>
</tr>
<tr>
<td>15</td>
<td>167.1 ± 5.4</td>
<td>75.8 ± 5.1</td>
<td>193.2 ± 7.3</td>
<td>54.6 ± 2.1</td>
</tr>
</tbody>
</table>
that the outflow tract was contoured as part of the LV cavity as recommended in the SCMR guidelines [11].

The STAPLE method was performed on each slice and each contour type independently, without taking into account any information about the geometry or anatomy of the myocardium or any pixel intensity values. The STAPLE method is not a vote counting mechanism, and therefore the functional consensus is not a simple average of mass or volume across readers. An example of the difference between voting and STAPLE consensus contours is shown in Fig. 5, demonstrating that STAPLE found an acceptable epicardial apex contour whereas voting did not.

Reader assessment

In this study, we developed a resource CMR dataset, where myocardial contours were defined through a consensus of seven independent expert readers. The dataset will be useful for training and assessing new readers on contouring CMR images. Feedback to new readers can include graphical displays of areas of maximum disagreement (Fig. 4). Quantitative measurements can be measured in terms of mean, maximum and standard deviation of the distance from the new contour to the consensus.

Differences in mass and volumes derived by new reader contours can be compared with Table 2. The CMR data are available on request from the Cardiac Atlas Project for the purpose of training new readers and benchmarking automated processing methods. Note the consensus is a fusion of a number of different core labs, and does not reflect the practice of any particular lab.

Limitations

Only fifteen cases of varying pathology were included in the study, due to the time consuming nature of manual contouring. However, the purpose of the study was to provide a resource for assessing automated methods and facilitating training. We therefore chose increased depth of readers from different centres over breadth of subjects. This provides considerable power for analysis of differences [15, 16]. Although resources are becoming available with many hundreds and even thousands of studies [5, 17], these provide manual contours from a small number (typically one or two) of centers. Our study therefore provides a unique resource representing the largest number of expert readers to date.

The SCMR guidelines [11] recommend either inclusion or exclusion of papillary muscles. Our study excluded papillary muscles from the LV mass, but papillary muscles are myocardial tissue, which ideally should be included. The formation of a manual ground truth from many expert readers would be very time consuming, but very valuable for validating automated methods [18, 19]. In the future, more widely automated methods are likely to increase the number of centers quantifying papillary mass. Right ventricular and atria consensus contours would also be useful to establish in the future.

Conclusion

We have estimated a set of consensus myocardial contours from SSFP CMR short-axis slices from expert readers representing experienced core labs around the world. The consensus contours achieved better agreement in LV mass and volumes than between readers.
**Fig. 4** An example of the automatically generated report visualizing the similarity between a reader (green) and consensus (red) contours. Magenta bands indicate the range of contour positions for all seven readers. Arrows indicate points on the reader contour with distance > 3 mm from the consensus contour. Top: endocardial; bottom: epicardial; left: ED; right: ES

**Fig. 5** Left: epicardial apical ED contours drawn by readers. Right: the difference of consensus between pixel voting (cyan contour) and the STAPLE algorithm (red contour)
This consensus dataset is valuable resource for the assessment of new readers, as a basis for multi-center analyses, as well as benchmarking automated methods. The main source of bias was small but consistent differences in contour placement in all areas.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
EN and AAY conceived the study and participated in the design of the study. DAB, MGF, CMK, SP, JS-M, JMW participated in data collection and in the design of the study, interpretation of results and drafting of the manuscript. All authors read and approved the final manuscript.

Acknowledgements
This project was supported by Awards R01HL087773 and R01HL121754 from the National Heart, Lung, and Blood Institute. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Heart, Lung, and Blood Institute or the National Institutes of Health.

Author details
1Department of Anatomy with Radiology, University of Auckland, Auckland, New Zealand. 2National Institute of Biomedical Imaging and Bioengineering, Bethesda, MD, USA. 3Departments of Medicine and Diagnostic Radiology, McGill University, Montreal, Canada. 4University of Virginia, Charlottesville, VA, USA. 5Harvard Medical School, Boston, USA. 6University of Leeds, Leeds, UK. 7Leiden University Medical Center, Leiden, Netherlands. 8Chanté University Medicine Berlin and HELIOS Klinikum Berlin-Buch, Berlin, Germany. 9Institute for Experimental and Translational Cardiovascular Imaging, DZHK Centre for Cardiovascular Imaging, University Hospital Frankfurt/Main, Frankfurt, Germany.

Received: 23 March 2015 Accepted: 1 July 2015
Published online: 28 July 2015

References