Mitral regurgitation quantification by cardiac magnetic resonance imaging (MRI) remains reproducible between software solutions [version 1; peer review: awaiting peer review]

Ciaran Grafton-Clarke, George Thornton, Benjamin Fidock, Gareth Archer, Rod Hose, Rob J. van der Geest, Liang Zhong, Andrew J. Swift, James M. Wild, Estefania De Gárate, Chiara Bucciarelli-Ducci, Sven Plein, Thomas A. Treibel, Marcus Flather, Vassilios S. Vassiliou, Pankaj Garg

1Medical School, University of East Anglia, Norwich, UK
2Institute for Cardiovascular Sciences, University College London Hospitals NHS Trust, London, UK
3Department of Infection, University of Sheffield, Sheffield, UK
4Department of Radiology, Leiden University Medical Center, Leiden, The Netherlands
5National Heart Centre, Duke NUS Graduate Medical School, Singapore, Singapore
6British Heart Institute, Britsol, UK
7Royal Brompton & Harefield NHS Trust, Guy's and St Thomas' NHS Foundation Trust, London, UK
8Leeds Institute of Cardiovascular and Metabolic Medicine, University of Leeds, Leeds, UK

Abstract

Background: The reproducibility of mitral regurgitation (MR) quantification by cardiovascular magnetic resonance (CMR) imaging using different software solutions remains unclear. This research aimed to investigate the reproducibility of MR quantification between two software solutions: MASS (version 2019 EXP, LUMC, Netherlands) and CAAS (version 5.2, Pie Medical Imaging).

Methods: CMR data of 35 patients with MR (12 primary MR, 13 mitral valve repair/replacement, and ten secondary MR) was used. Four methods of MR volume quantification were studied, including two 4D-flow CMR methods (MR\text{MVAV} and MR\text{Jet}) and two non-4D-flow techniques (MR\text{Standard} and MR\text{LVRV}). We conducted within-software and inter-software correlation and agreement analyses.

Results: All methods demonstrated significant correlation between the two software solutions: MR\text{Standard} (\(r=0.92, p<0.001\)), MR\text{LVRV} (\(r=0.95, p<0.001\)), MR\text{Jet} (\(r=0.86, p<0.001\)), and MR\text{MVAV} (\(r=0.91, p<0.001\)). Between CAAS and MASS, MR\text{Jet} and MR\text{MVAV} compared to each of the four methods, were the only methods not to be associated
with significant bias.

Conclusions: We conclude that 4D-flow CMR methods demonstrate equivalent reproducibility to non-4D-flow methods but greater levels of agreement between software solutions.

Keywords
Magnetic resonance imaging; Mitral valve insufficiency;
Reproducibility of results.

Corresponding author: Ciaran Grafton-Clarke (Ciarang-c@hotmail.com)

Author roles: Grafton-Clarke C: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Software, Validation, Writing – Original Draft Preparation, Writing – Review & Editing; Thornton G: Conceptualization, Data Curation, Validation, Writing – Original Draft Preparation, Writing – Review & Editing; Fidock B: Conceptualization, Data Curation, Methodology, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing; Archer G: Conceptualization, Data Curation, Formal Analysis, Methodology, Writing – Original Draft Preparation, Writing – Review & Editing; Hose R: Conceptualization, Data Curation, Investigation, Writing – Review & Editing; van der Geest RJ: Conceptualization, Investigation, Methodology, Resources, Software, Writing – Review & Editing; Zhong L: Conceptualization, Investigation, Methodology, Writing – Review & Editing; Swift AJ: Conceptualization, Data Curation, Methodology, Resources, Writing – Review & Editing; Wild JM: Conceptualization, Data Curation, Investigation, Methodology, Writing – Review & Editing; De Gárate E: Conceptualization, Data Curation, Investigation, Methodology, Writing – Review & Editing; Bucciarelli-Ducci C: Conceptualization, Data Curation, Investigation, Writing – Review & Editing; Plein S: Conceptualization, Project Administration, Supervision, Validation, Writing – Review & Editing; Treibel TA: Conceptualization, Data Curation, Methodology, Writing – Review & Editing; Flather M: Conceptualization, Data Curation, Supervision, Writing – Review & Editing; Vassiliou VS: Conceptualization, Data Curation, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing; Garg P: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Supervision, Validation, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: Dr P Garg is a clinical advisor for Pie Medical Imaging and Medis Medical Imaging.

Grant information: This work was supported by Wellcome Trust [220703, 215799]. For the purpose of open access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2021 Grafton-Clarke C et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Grafton-Clarke C, Thornton G, Fidock B et al. Mitral regurgitation quantification by cardiac magnetic resonance imaging (MRI) remains reproducible between software solutions [version 1; peer review: awaiting peer review] Wellcome Open Research 2021, 6:253 https://doi.org/10.12688/wellcomeopenres.17200.1

First published: 05 Oct 2021, 6:253 https://doi.org/10.12688/wellcomeopenres.17200.1
Introduction
Mitral regurgitation (MR) is one of the most common types of valvular heart disease and is one of the most frequent indications for valve surgery. Even though echocardiography remains the first-line investigation for MR assessment, recent evidence suggests that cardiovascular magnetic resonance (CMR) quantitative assessment of MR is more precise and has a better prognostic association. One of the key strengths of CMR quantification of MR is that it allows many different ways to quantify MR. These include direct and indirect methods using standard techniques and emerging four-dimensional (4D) flow methods.

Our recent work demonstrated that 4D-flow methods of MR quantification may offer superior precision for reproducibility compared to standard methods. In practice, a combination of standard and 4D-flow methods of MR quantification can be used to build confidence in reporting CMR images and clinical decision-making. Our previous work involved the use of a research software solution from Leiden lab (MASS). MASS is not currently a commercial software package for clinical use and is limited to research applications only. Moreover, there is a paucity of evidence evaluating the reproducibility of MR volume quantification between different software solutions across the breadth of methods. Demonstrating reproducibility between different software solutions is vital as clinical outcome research within CMR imaging is multiplatform and multicentre. It is essential that the data generated from analysis is accurate, precise, and reproducible, regardless of which software platform is used.

The primary objective of this research was to investigate the reproducibility and agreement in MR volume quantification between two software solutions (CAAS, version 5.2, Pie Medical Imaging) using subjects from previously published cohorts spanning the spectrum of MR disease states. Using CAAS, we also conducted within-software agreement analysis between different methods of MR volume quantification. Third, we present interobserver reproducibility analysis within CAAS across the four methods of MR volume quantification.

Methods
Study population
The subjects included within this study have been reported on in other published works. In brief, the data relates to a UK multicentre prospective study involving 35 subjects with MR diagnosed on echocardiography. Recruited from outpatient cardiology clinics at two centres with dedicated mitral valve services (Sheffield and Leeds) between January 2015 – December 2020, 12 subjects had primary MR, ten subjects had secondary MR, and 13 subjects had mitral valve replacement (MVR). Patients with significant valvular stenosis and cardiac shunts were not considered eligible.

Ethics
This study was approved by the National Research Ethics Committee in the UK (17/LO/0283 and 12/YH/0169). Informed written consent was obtained from all subjects before participation.

CMR protocol
At Sheffield, CMR was performed on a 3.0 Tesla Philips Healthcare system (Achieva TX) equipped with a 28-channel coil and Philips dStream digital broadband MR architecture technology. In Leeds, CMR was performed on a 1.5 Tesla Philips Healthcare system (Ingenia).

The CMR protocol included baseline surveys, cines (vertical long-axis, horizontal long-axis, short-axis contiguous left ventricle volume stack, 3-chamber, and aortic root) and 4D-flow acquisition. Cine images were acquired during end-expiratory breath-holds with a balanced steady-state free precession, single-slice breath-hold sequence. Procedures relating to 4D-flow pre-processing were delivered in accordance with established standards of practice.

Image analysis
Image analysis was completed within two CMR software solutions: MASS software (version 2019 EXP, LUMC, Netherlands) and CAAS MR Solutions (version 5.2). The image analysis and MR quantification methods for the MASS platform are published elsewhere. In CAAS, both aliasing correction and phase offset correction were applied.

In total, four quantification methods for MR were computed within the CAAS platform, aligning with the methods used within MASS (Figure 1). One assessor completed the analysis of all 35 subjects within CAAS, blinded to the data generated from MASS.

1. $MR_{\text{Standard}} (LVSV - AoPC)$
   Left ventricular stroke volume (LVSV) was determined through endocardial segmentation of the short-axis cine stack. Aortic stroke volume was obtained using a static reformatted aortic phase-contrast (AoPC) plane through the sino-tubular junction.

2. $MR_{\text{LVSV}} (LVSV - RVSV)$
   Right ventricular stroke volume (RVSV) was determined by segmentation of the RV in the short-axis cine stack. This method was not used in the ten patients with secondary MR, given the regular presence of concurrent MR and tricuspid regurgitation.

3. $MR_{\text{MVR}} (4D$-flow mitral forward flow - 4D-flow aortic forward flow)$
   Using retrospective mitral valve and aortic valve tracking within the four-chamber cine and three-chamber cine, respectively, a phase-contrast, valvular formatted plane was generated. Using the formatted valvular plane, we segmented the forward
flow whilst taking into account the through-plane motion of the valve plane.

4. **MR** (4D-flow direct jet assessment)

Jets of MR were directly quantified from the 4D-flow dataset. The jet(s) were first identified in multiple long-axis chamber views. Where available, the four-chamber view was used to draw a reformatted plane perpendicular to the regurgitant jet within the left atrium for each phase it was present. If multiple, jet volumes were summated to provide a total MR volume.

**Interobserver reproducibility**

Interobserver tests were performed by two investigators (CGC, PG) blinded to the results of each other. A random mix of ten subjects was studied, where each investigator estimated MR volume using the four methods previously described. Each observer had at least two years of CMR experience.

**Statistical analysis**

All statistical analysis were completed using **SPSS** version 25, though Microsoft Excel could also be used. All continuous parameters are reported as mean ± standard deviation (SD).

Statistical parameters to assess inter-software and within-software MR quantification method correlation were calculated using Pearson correlation coefficient. Agreement between methods of MR quantification within-software (CAAS) and between software’s (CAAS versus MASS) was calculated using Bland-Altman statistics where the mean difference between two methods was reported as the relative risk of bias (measured in ml). For all analyses, p < 0.05 was deemed to be statistically significant. Defined *a priori*, bias between methods of greater than 5 ml was felt to be clinically significant, as determined through consensus amongst study investigators.

**Results**

All 35 subjects underwent CMR imaging, and at the time of investigation, were in sinus rhythm. Demographic and clinical data for the 35 subjects are presented in Table 1. As quantified using CAAS, the average MR volume (across all four methods) for subjects with primary MR was 30.5 ml, 16.4 ml for subjects with secondary MR, and 3.2 ml in those with a replaced/repaired mitral valve.

**Inter-software correlation and agreement**

Quantification of MR in CAAS correlated strongly with the values from MASS for all four methods of assessment (Table 2). **MR** was the most strongly correlated method between...
Despite being the most strongly correlated method between software solutions, MR\textsubscript{Standard} was the only method to result in significant bias in agreement between CAAS and MASS MR quantification (bias 2.7 ml, p = 0.045) (Figure 2). The degree of bias for the other methods was 2.3 ml for MR\textsubscript{MVAV} (p = 0.137), 1.4 ml for MR\textsubscript{LVRV} (p = 0.338) and -2.5 ml for MR\textsubscript{Jet} (p = 0.169). Of note, when we performed subgroup analysis of agreement for the MR\textsubscript{Standard} stratified by MR type, it was identified that MR\textsubscript{Standard}, when used for MR quantification in

### Table 1. Study participant demographics and clinical data.

<table>
<thead>
<tr>
<th></th>
<th>Primary MR</th>
<th>Secondary MR</th>
<th>MVR</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of subjects</td>
<td>12</td>
<td>10</td>
<td>13</td>
<td>-</td>
</tr>
<tr>
<td>Age (years)</td>
<td>67 ± 11</td>
<td>68 ± 11</td>
<td>62 ± 11</td>
<td>0.97</td>
</tr>
<tr>
<td>Male, n (%)</td>
<td>6 (50.0)</td>
<td>6 (60.0)</td>
<td>13 (100.0)</td>
<td>0.03</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>167 ± 8</td>
<td>167 ± 9</td>
<td>177 ± 6</td>
<td>0.04</td>
</tr>
<tr>
<td>Weight (Kg)</td>
<td>75 ± 11</td>
<td>77 ± 11</td>
<td>93 ± 19</td>
<td>0.01</td>
</tr>
<tr>
<td>Diabetes mellitus (n)</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0.87</td>
</tr>
<tr>
<td>Smoker (n)</td>
<td>7</td>
<td>7</td>
<td>5</td>
<td>0.98</td>
</tr>
<tr>
<td>Atrial fibrillation (n)</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>0.26</td>
</tr>
<tr>
<td>Ischaemic heart disease (n)</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>-</td>
</tr>
<tr>
<td>NYHA class</td>
<td>2.4 ± 0.9</td>
<td>1.3 ± 0.6</td>
<td>1.7 ± 0.7</td>
<td>0.01</td>
</tr>
</tbody>
</table>

MR=magnetic resonance; MRV-mitral valve replacement; NYHA-New York Heart Association.

### Table 2. Correlation and agreement analysis between CAAS and MASS mitral regurgitation quantification methods.

Correlation analysis using the Pearson correlation coefficient (denoted Correlation) and agreement analysis using Bland-Altman statistics (denoted Bias). The table provides within-vendor analysis (i.e., correlation and agreement between each method within CAAS software solutions) and inter-vendor analysis (correlation and agreement for each method between CAAS and MASS software solution). For agreement analysis, bias refers to the mean difference between two methods of MR volume quantification (measured in ml) and is deemed statistically significant if the corresponding p-value (denoted P) is < 0.05. For negative bias values, this indicates that the method used in CAAS (uppermost panel) to quantify MR is systematically lower than the method in either CAAS (for within-vendor analysis) or MASS (for inter-vendor analysis). For correlation analysis, a p-value < 0.05 is deemed statistically significant. MR=magnetic resonance.

<table>
<thead>
<tr>
<th></th>
<th>CAAS</th>
<th>MASS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MR\textsubscript{Standard}</td>
<td>MR\textsubscript{LVRV}</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correlation</td>
<td>R</td>
<td>P</td>
</tr>
<tr>
<td>Bias</td>
<td>-</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>P</td>
<td>0.73</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td>P</td>
<td>0.78</td>
</tr>
</tbody>
</table>

Software solutions (R 0.95, p < 0.001), followed by MR\textsubscript{Standard} (r = 0.92, p < 0.001) and MR\textsubscript{MVAV} (r = 0.91, p < 0.001). MR\textsubscript{Jet} (r = 0.86, p < 0.001) was the least strongly correlated method.
Figure 2. (A) Bland-Altman plots for mitral regurgitation within CAAS. Each plot represents a comparison between two methods within CAAS. Bias refers to the mean difference between the methods of mitral regurgitation volume quantification (measured in ml) and is deemed statistically significant (i.e., high risk of systematic bias) if the corresponding p-value is < 0.05. (B) Bland-Altman plots for MR quantification between CAAS and MASS. Each plot represents a comparison between like-for-like methods of mitral regurgitation volume quantification between the two software solutions. MR-magnetic resonance.
subjects with MVR, demonstrated poor levels of agreement (bias 6.7 ml, p = 0.007). This contrasts with the agreement in subjects with primary and secondary MR, where MR\textsubscript{\text{Standard}} was associated with low bias (-1.0 ml, p = 0.604 and 2.1 ml, p = 0.392, respectively). With specific reference to the 4D-flow methods of MR quantification, MR\textsubscript{MVAV} and MR\textsubscript{Jet}, there was excellent correlation and low bias between these methods in CAAS and all four methods within MASS.

Within-software correlation and agreement
Using CAAS, we compared each method to each other to determine correlation and agreement/bias. All methods were strongly correlated to each other, with r coefficients ranging from 0.73 to 0.91. Of all method comparisons, MR\textsubscript{\text{Standard}} and MR\textsubscript{LVRV} were the most positively correlated (r = 0.91 p < 0.001). The least strongly correlated methods were MR\textsubscript{\text{Standard}} and MR\textsubscript{Jet} (r = 0.73, p < 0.001) and also MR\textsubscript{LVRV} and MR\textsubscript{Jet} (r = 0.76, p < 0.001).

Despite being the most strongly correlated, MR\textsubscript{\text{Standard}} and MR\textsubscript{LVRV}, when compared to each other, were associated with significant levels of agreement bias (6.2 ml, p = 0.009). Further to this, MR\textsubscript{LVRV} had significant bias when compared to both MR\textsubscript{MVAV} (5.2 ml, p = 0.020) and MR\textsubscript{Jet} (7.2, p = 0.007). Aside from comparisons with MR\textsubscript{LVRV}, the 4D-flow methods of MR quantification were associated with low levels of bias when compared to each other and to MR\textsubscript{\text{Standard}}.

Interobserver reproducibility
Reproducibility in analysis between two independent assessors with CAAS demonstrated excellent agreement across all four methods of MR volume quantification (Table 3). The 4D-flow methods of quantification were the most strongly correlated between observers (MR\textsubscript{\text{Jet}}, r = 0.99, p < 0.001; MR\textsubscript{MVAV} 0.98, p < 0.001). MR\textsubscript{\text{Standard}} and MR\textsubscript{LVRV} were also strongly correlated (0.96 and 0.94, respectively, p < 0.001). Only MR volume quantification using the MR\textsubscript{Jet} method between two observers demonstrated significant bias\(^5\). MR\textsubscript{\text{Standard}} and MR\textsubscript{LVRV} and MR\textsubscript{MVAV} methods of quantification were not significantly biased between two observers.

Discussion
We have demonstrated that quantification of mitral regurgitation is consistent between two different software solutions. We have also demonstrated that within the CAAS platform, there are high levels of agreement between all methods of quantification. Between software solutions, MR\textsubscript{\text{Standard}} was the only method to result in significant bias and was identified to be due to subjects with mitral valve replacement. We speculate this may be due to the challenges in segmenting the short-axis basal slices in subjects with a MVR. Of note, despite the bias associated with the MR\textsubscript{\text{Standard}} method being determined as statistically significant, the quantity of MR volume of 2.7 ml is not clinically significant.

Between methods in CAAS, the degree of correlation between all methods was excellent. The MR\textsubscript{\text{Standard}} method was not only strongly correlated with the MR volume quantification methods utilising 4D-flow techniques, but there was a low risk of bias between MR\textsubscript{\text{Standard}} and both MR\textsubscript{MVAV} and MR\textsubscript{Jet} methods of quantification. We have therefore demonstrated that within CAAS, with reference to the MR\textsubscript{\text{Standard}} method, agreement is best demonstrated with 4D-flow techniques. We have also shown that between the two software platforms, MR volume quantification using the 4D-flow techniques, is both highly reproducible, and is not associated with significant bias, which was not the case for the non-4D-flow techniques.

A previous multicentre study demonstrated that automated valve tracking on CAAS can provide consistent valvular flow quantification\(^6\). Our study complements their work and demonstrates interoperability between different CMR methods of MR quantification. This becomes critically important in routine clinical practice for increasing the confidence of reporting MR severity. In addition, in this study, we have demonstrated agreement and consistency in MR quantification between two software solutions. This is important for the clinical translation of all the methods of MR quantification by CMR described in our study.

Our previous work demonstrated that 4D-flow methods of MR quantification, in particular MR\textsubscript{MVAV}, is superior to other methods of MR quantification for reproducibility as it enhances precision\(^7\). As research involving 4D-flow CMR techniques continues to gain interest, there is an evolving need for large multicentre studies with clinical outcomes to provide answers to key clinical questions. It is therefore essential for the research and clinical communities to have confidence that regardless of the software platform used for analysis, the data output is comparable between platforms and can confidently be combined without risk of significant bias.

This study has several limitations. Firstly, patients with MVR and secondary MR only had mild to moderate MR. Secondly, due to lower MR volume in MVR and secondary MR cases,

Table 3. Interobserver reproducibility analysis.

<table>
<thead>
<tr>
<th></th>
<th>Pearson Correlation</th>
<th>Bland Altman</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>r</td>
<td>p-value</td>
</tr>
<tr>
<td>MR\textsubscript{\text{Standard}}</td>
<td>0.964</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>MR\textsubscript{LVRV}</td>
<td>0.939</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>MR\textsubscript{MVAV}</td>
<td>0.980</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>MR\textsubscript{Jet}</td>
<td>0.988</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Correlation analysis using the Pearson correlation coefficient (denoted Correlation) and agreement analysis using Bland-Altman statistics (denoted Bias) between two observers within the CAAS software solution. MR-magnetic resonance.
the relative bias may appear larger in Bland-Altman analysis. Thirdly, we have only used one commercially available CMR software for comparison. And finally, this study did not evaluate inter-scan variability in MR volume quantification.

**Data availability**

**Underlying data**

Harvard Dataverse: Mitral regurgitation quantification by cardiac MRI between software solutions. https://doi.org/10.7910/DVN/I8S00H.

This project contains the following underlying data:

- Data Upload.tab (demographic data; functional data and outputted 4D-flow data from both software solutions; inter-observer data between assessor 1 and 2)

Data are available under the terms of the Creative Commons Zero “No rights reserved” data waiver (CC0 1.0 Public domain dedication).

Raw CMR images were not uploaded in order to protect the identity of the subjects. Access can be requested by contacting the corresponding author (Ciarang-c@hotmail.com). Access to the raw CMR images will be granted for the purpose of re-analysis relating to the primary aims of this research.

**Consent**

Written informed consent for publication of the participants’ data and data resulting from analysis of their cardiac imaging was obtained from the participants.

**References**


