Comparison of (semi-)automatic and manually adjusted measurements of left ventricular function in dual source computed tomography using three different software tools

G. J. de Jonge · P. M. A. van Ooijen · J. Overbosch · A. Litcheva Gueorguieva · M. C. Janssen-van der Weide · M. Oudkerk

Received: 21 May 2010 / Accepted: 5 October 2010 / Published online: 23 October 2010 © The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract To assess the accuracy of (semi-)automatic measurements of left ventricular (LV) functional parameters in cardiac dual-source computed tomography (DSCT) compared to manually adjusted measurements in three different workstations. Forty patients, who underwent cardiac DSCT, were included (31 men, mean age 58 ± 14 years). Multiphase reconstructions were made with ten series at every 10% of the RR-interval. LV function analysis was performed on three different, commercially available workstations. On all three workstations, end-systolic volume (ESV), end-diastolic volume (EDV), LV ejection fraction (LVEF) and myocardial mass (MM) were calculated as automatically as possible. With the same DSCT datasets, LV functional parameters were also calculated with as many manual adjustments as needed for accurate assessment for all three software tools. For both semi-automatic as well as manual methods, time needed for evaluation was recorded. Paired t-tests were employed to calculate differences in LV functional parameters. Repeated measurements were performed to determine intra-observer and inter-observer variability. (Semi-)automatic measurements revealed a good correlation with manually adjusted measurements for Vitrea (LVEF \( r = 0.93 \), EDV \( r = 0.94 \), ESV \( r = 0.98 \) and MM \( r = 0.94 \)) and Aquarius (LVEF \( r = 0.96 \), EDV \( r = 0.94 \), ESV \( r = 0.98 \) and MM \( r = 0.96 \)). Also, good correlation was obtained for Circulation, except for LVEF (LVEF \( r = 0.45 \), EDV \( r = 0.93 \), ESV \( r = 0.92 \) and MM \( r = 0.86 \)). However, statistically significant differences were found between (semi-)automatically and manually adjusted measurements for LVEF (\( P < 0.05 \)) and ESV (\( P < 0.001 \)) in Vitrea, all LV functional parameters in Circulation (\( P < 0.001 \)) and EDV, ESV and MM (\( <0.001 \)) in Aquarius Workstation. (Semi-)automatic measurement of LV functional parameters is feasible, but significant differences were found for at least two different functional parameters in all three workstations. Therefore, expert manual correction is recommended at all times.

Keywords Dual source computed tomography · Left ventricular function · Software comparison · Heart · Function · Computed tomography

Introduction

Left ventricular (LV) functional parameters, especially LV ejection fraction (LVEF) and mass, are
well-known predictors of morbidity and mortality in patients with ischemic heart disease [1]. Therefore, accurate and reliable quantification of these parameters is important for clinical diagnosis and the prediction of prognosis in patients with suspected or known coronary artery disease [2, 3]. Assessment of LV function can be performed by various non-invasive imaging modalities as well as by invasive contrast ventriculography. Currently, short-axis magnetic resonance imaging (MRI) is considered to be the reference standard for LV function analysis, due to its high temporal resolution and superior tissue contrast between blood pool and myocardium [4–7]. Cardiovascular MRI has been shown to be highly accurate and reproducible for the assessment of cardiac function and mass [8]. In recent years, retrospective ECG-gated multi-detector computed tomography (MDCT) has developed a role in cardiac imaging and is increasingly being used in clinical practice. In many studies the image quality of thin-section MDCT coronary angiograms have been investigated and a high diagnostic accuracy with good sensitivity and specificity values has been reported [9–11]. Every cardiac CT dataset with retrospective ECG-gating through the whole cardiac cycle also supplies the information needed to determine LV functional parameters. A good agreement of LV functional parameters with cardiac MDCT compared to other known non-invasive modalities, including cine MRI, was shown previously [4, 5, 12, 13]. To evaluate cardiac function with MDCT datasets, many different software tools are commercially available nowadays that rely on the segmentation of the left ventricular blood pool in end-systole and end-diastole for determination of the LV functional parameters. Manual reconstruction of short-axis images from raw MDCT datasets with axial image reconstruction and the manual drawing of endo- and epicardial contours can be very time consuming [14, 15]. Therefore, semi-automatic software tools have been developed using a region-growing segmentation approach, based on the value of the Hounsfield Units (HU) within the LV cavity. The region-growing software algorithm has been reported to be significantly faster compared to the traditional LV function quantification with the Simpson method [16]. On the other hand, the threshold-based segmentation software can also be extensively manipulated manually, which can be time-consuming and makes it operator-dependable. In this study, the (semi-)automatic and manually adjusted LV functional measurements derived from retrospective ECG-gated dual-source CT datasets were compared using three different threshold-based cardiac CT analysis software packages. The aim of this study is to determine if (semi-)automatic measurements of LV functional parameters in cardiac dual-source computed tomography are as accurate as manually adjusted measurements.

**Materials and methods**

**Patients**

The study population consisted of forty consecutive patients (31 men; mean age 58 years ± 14, range 26–86 years) who all underwent cardiac CT angiography in our hospital. Reasons for cardiac DSCT were evaluation of coronary bypass graft patency (n = 4), exclusion of aberrant coronary anatomy (n = 11), exclusion of coronary stenosis (n = 8) or evaluation of the cardiac anatomy (n = 4). Thirteen patients participated in two different research protocols: a prospective study on the comparison of DSCT coronary angiography and conventional coronary angiography (CAG) (n = 9) and a screening study for patients with peripheral arterial disease (n = 4). Nine patients had experienced previous myocardial infarction. Exclusion criteria for this study were renal insufficiency (creatinine > 120 μmol/l), hyperthyreoidism and contrast allergy. No beta-blockers were administered prior to the scan to lower the heart rate.

**DSCT data acquisition**

All cardiac CT angiograms were performed on a DSCT system (Somatom Definition, Siemens Medical Solutions, Forcheim, Germany). Retrospective ECG-triggered images were obtained in the cranio-caudal direction starting at the tracheal bifurcation to well below the apex, during a single breath hold. The following scan parameters were used: tube voltage was 120 kV and tube current 410 mAs/rot for both tubes. Scanning was performed with a 64 × 0.6-mm collimation. The gantry rotation time was 330 ms and the pitch ranged from 0.2 to 0.5 s, adapted automatically to the heart rate. The ECG pulsing window was
set to 20–70% of the RR-interval for all patients. Coronary and left ventricle enhancement was achieved by intravenous injection of 80 ml of non-ionic contrast agent (Iomeprol 400 mg/ml, Iomeron® 400, Bracco, Italy) via an antecubal vein at a flow rate of 5 ml/s followed by a saline bolus flush. Bolus timing was performed automatically with a region of interest set in the descending aorta, with a fixed threshold of 100 HU. For left ventricular function analysis, reconstructions were made at every 10% of the RR-interval with 2.0 mm slice thickness. A monosegmental reconstruction algorithm was used for all image-reconstructions. The reconstruction kernel was set to a B26f Heartview smooth kernel.

DSCT data analysis

All forty multiphase datasets were transferred to three different cardiac workstations: Circulation version VE22A (Siemens Medical Systems, Germany), Vitrea®2 version 4.0.0.0 (Vital Images, Minnetonka, MN) and Aquarius Workstation version 3.5.2.1 (TeraRecon, San Mateo, CA). Two readers assessed LV functional parameters semi-automatic as well as manually adjusted. Length of time needed for complete on all software tools, both automatically and manually adjusted, was recorded for both readers. All three software tools use a threshold-based volume measurement algorithm, which quantifies the volume of voxels with densities between two specific Hounsfield unit values. Thus, papillary muscles and trabeculae are excluded from the LV blood pool volume.

To obtain LV functional parameters as automatically as possible for Circulation workstation, only one mouse-click is needed in the majority of patients. End-diastolic (ED) and end-systolic (ES) phases are in each patient automatically set by the software at 20 and 80% of the RR-interval. Then, the left ventricular volumes are automatically segmented based on Hounsfield units and a point between the anteroseptal and anterior position has to be marked. After this, the endo- and epicardial contours are automatically traced and LV functional parameters are displayed. In a few patients, the LV cavity was not automatically segmented. In such cases, the mitral valve plane had to be placed manually. After this, the LV volume could be segmented by clicking at a random point in the contrast-enhanced LV blood pool. When calculating LV parameters with manual adjustments, there are several options to change if necessary: the ED and ES phases can be set manually based on the visually largest and smallest LV volume, the mitral valve plane can be manually set and endo- and epicardial contours can be manually altered.

Vitrea®2 software automatically calculates and displays the functional measurements when a multiphase CT dataset is loaded into the workstation, so no user interaction is required. But if necessary, several adjustments can be made: the mitral valve plane and apex plane can be changed and endo- and epicardial contours can be manually altered.

In Aquarius workstation, the only manual interaction which is required to obtain LV functional parameters is determining the mitral valve plane in horizontal and vertical long-axis planes. The options for manual adjustments are changing the threshold of Hounsfield units for segmentation of the LV cavity and the endo- and epicardial contours can be changed.

In all patients, images analysis was performed by two readers (both with 2 years experience in cardiac imaging) to assess the interobserver variability. Image analysis was repeated in a random order by one reader to determine intraobserver variability. The observers were had globally an equal experience with the three different software tools.

Statistical analysis

All data are summarized as mean ± standard deviation (SD). A paired t-test was performed to calculate differences in LV functional parameters between automatic and manual results for all software tools. Bland–Altman plots were made to show mean difference and limits of agreement between automatic and manual results [17]. A P-value < 0.05 was considered statistically significant. Pearson’s correlation coefficient was calculated to assess the correlation between the automatic and manually adjusted measurements. The reproducibility of the LV functional parameters was evaluated by calculating the intra-observer and inter-observer variability. The variability was defined as the absolute difference between the corresponding repeated measurements expressed in percent of their mean. Variability values obtained for each parameter in each patient for each software package were averaged over the whole group of patients and represented as mean ± SD.
Results

All cardiac DSCT examinations were successfully performed without adverse events. The calculated mean effective radiation dose based on the used scan protocol was 7.3 mSv [18]. The average heart rate during DSCT examination was 66.1 ± 13.8 beats/min (range 46–104 beats/min). All studies were of adequate image quality to evaluate LV functional parameters. The mean values of LVEF, EDV, ESV and MM for automatically as well as manually adjusted measurements are represented in Table 1 for the three different workstations. Statistically significant differences were found between (semi-)automatic and manually adjusted measurements for LVEF (P < 0.05) and ESV (P < 0.001) in Vitrea, all LV functional parameters in Circulation (P < 0.001) and EDV, ESV and MM (<0.001) in Aquarius Workstation.

(Semi-)automatic measurements revealed a good correlation with manually adjusted measurements for Vitrea (LVEF r = 0.93, EDV r = 0.94, ESV r = 0.98 and MM r = 0.94) and Aquarius (LVEF r = 0.96, EDV r = 0.94, ESV r = 0.98 and MM r = 0.96). Good correlation was obtained for Circulation as well, except for LVEF (LVEF r = 0.45, EDV r = 0.93, ESV r = 0.92 and MM r = 0.86) (Table 1).

Bland–Altman plots showed a good agreement among automatically and manually adjusted measurements for Vitrea (mean difference LVEF −1.6 ± 4.4, EDV 2.7 ± 16.4, ESV 3.7 ± 6.4, and MM 0.8 ± 19.1) and Aquarius (mean difference LVEF −0.3 ± 3.5, EDV −18.0 ± 17.2, ESV −7.9 ± 11.5, MM 28.7 ± 23.7). Agreement for Circulation was moderate (LVEF −26.1 ± 13.0, EDV −24.2 ± 19.1, ESV 28.1 ± 18.8, and MM 48.3 ± 44.0) (Fig. 1a–c). The Bland–Altman plots show several outliers, with the highest mean difference of 13.0% (Vitrea), 44.5% (Circulation) and 11.8% (Aquarius) for LVEF, 67.5 ml (Vitrea), 94.5 ml (Circulation) and 96.8 ml (Aquarius) for EDV, 12.5 ml (Vitrea), 37.5 ml (Circulation) and 63.5 ml (Aquarius) for ESV and 66.5 g (Vitrea), 128.5 g (Circulation), 85.7 g (Aquarius) for MM (Table 2).

For all workstations, automatic measurements were significantly faster compared to manually adjusted measurements [Vitrea 17.6 ± 3.2 vs. 417.7 ± 91.0 s (P < 0.001) Circulation 61.3 ± 20.4 vs. 352.8 ± 67.5 s (P < 0.001) and Aquarius 74.4 ± 8.9 vs. 213.9 ± 32.6 s (P < 0.001)]. Intra- and interobserver variability were good (Tables 3, 4).

Discussion

In this study, we examined whether automatic measurements for LV functional parameters with DSCT data are as accurate as measurements with manual adjustments. Three different, (semi-)automatic software tools were used; these software tools all use a region-growing algorithm, which segments the LV ventricle based on the HU value from apex to mitral valve. Traditionally, LV functional parameters are calculated based on manually drawn contours in short axis slices according to the Simpson method [19]. Several studies demonstrate an excellent reliability of manual contour tracing with the modified Simpson method for calculating LV functional parameters [14, 19, 20]. The main drawback for this method is the time needed for evaluation [20].
Fig. 1 Results of Bland–Altman plots for LVEF, EDV, ESV and MM between (semi-)automatic and manually adjusted measurements for the three different workstations.
Therefore, many commercial (semi-)automatic software packages with a region-growing segmentation approach have been developed which require very minimal user interactions to shorten the examination time for the clinician. To obtain LV functional parameters, minimal user interaction is needed in these packages. However, even in these (semi-)automatic software tools LV functional measurements can show variability.

### Table 2: Maximum mean difference in measurements between automatic and manually adjusted measurements

| Software   | LVEF (%) | EDV (ml) | ESV (ml) | MM (g) |
|------------|----------|----------|----------|--------|
| Vitrea     | 13.0     | 67.5     | 12.5     | 66.5   |
| Circulation| 44.5     | 94.5     | 37.5     | 128.5  |
| Aquarius   | 11.8     | 96.8     | 63.5     | 85.7   |

### Table 3: Intra-observer variability

| Software   | LVEF | EDV   | ESV   | MM   |
|------------|------|-------|-------|------|
| Vitrea auto| 0.0  | 0.0   | 0.0   | 0.0  |
| Vitrea manual| 1.7 | 1.8   | 4.1   | 2.1  |
| Circulation auto| 4.6 | 1.4   | 1.1   | 2.9  |
| Circulation manual| 2.4 | 2.0   | 2.2   | 4.0  |
| Aquarius auto| 2.3 | 5.1   | 5.0   | 6.4  |
| Aquarius manual| 1.5 | 2.8   | 3.5   | 4.4  |

The variability is defined as the absolute difference between the corresponding repeated measurements expressed in percentage of their mean.

### Table 4: Inter-observer variability

| Software   | LVEF   | EDV   | ESV   | MM   |
|------------|--------|-------|-------|------|
| Vitrea auto| 0.0    | 0.0   | 0.0   | 0.0  |
| Vitrea manual| 4.8  | 4.3   | 9.8   | 4.3  |
| Circulation auto| 2.5  | 1.1   | 0.1   | 3.5  |
| Circulation manual| 7.1  | 8.7   | 4.7   | 7.4  |
| Aquarius auto| 3.4    | 8.4   | 9.1   | 8.7  |
| Aquarius manual| 3.3  | 5.5   | 9.8   | 9.6  |

The variability is defined as the absolute difference between the corresponding repeated measurements expressed in percentage of their mean.

Therefore, many commercial (semi-)automatic software packages with a region-growing segmentation approach have been developed which require very minimal user interactions to shorten the examination time for the clinician. To obtain LV functional parameters, minimal user interaction is needed in these packages. However, even in these (semi-)automatic software tools LV functional measurements can show variability.
be manipulated extensively. In all software tools, endo- and epicardial contours can be changed, as well as the mitral valve plane (and apex plane for Vitrea). In Circulation, ED and ES phases can be manually altered, whereas the other two workstations calculate the largest and smallest LV volumes automatically.

First, a good correlation was found for all parameters between (semi-)automatic and manually adjusted parameters for both Vitrea and Aquarius. The Bland–Altman plots showed a good agreement with acceptable limits of agreement. On the other hand, for these two workstations, significant differences were found using paired t-tests for LVEF and ESV in Vitrea and EDV, ESV and MM in Aquarius. Also, several outliers in the Bland–Altman plot can be observed with quite large differences between automatically and manually adjusted measurements (as high as 13% difference for LVEF, 96.8 ml for EDV, 63.5 ml for ESV and 85.7 g for MM, see Table 4) which is not acceptable in clinical practice. This implies that, although in most patients the differences are small (as can be observed in the Bland–Altman plots) automatic measurements are not interchangeable per definition. It suggests that visual control of the (semi-)automatic segmentation is still mandatory. This was also concluded by Mühlenbruch et al. [21] who compared a region-growing software algorithm with manually drawn contours in short axis slices. Proper segmentation by the region-growing algorithm was feasible in 65% of patients. Another study reports a failure of 6 out of 26 patients to successfully segment the left ventricular blood pool by a region-growing algorithm in a comparison study with clinical echocardiography [22]. A study comparing LV functional parameters obtained with different (semi-)automatic region-growing software tools, also shows a large difference in LVEF in individual patients when automatic measurements are compared between the software tools [23]. Difficulties with segmentation were reported to be caused by pacemaker leads and ventricular septal defects [22] and low signal-to-noise ratio or poorer contrast enhancement of the left ventricle [21]. In our study, in a few patients the mitral valve plane was not detected correctly but at a higher or lower level. A possible cause for this could be the inability of the software to detect the valve when opened.

Secondly, a good correlation was observed for EDV, ESV and MM in Circulation but correlation for LVEF was 0.45. The Bland–Altman plots showed large mean differences and wide limits of agreement for all LV functional parameters, and t-tests demonstrated significant differences for all parameters. These results can be explained by the fact that Circulation is the only software tool, among the three software tools used, which requires manual setting of ED and ES phases. When this is not done manually, the software assumes the fixed setting of 20 and 80% of the RR-interval. Therefore, the results are inferior to the other two software tools, which calculate ES and ES phases automatically based on the automatically determined ventricular volumes. It is likely that the results between (semi-)automatic and manually adjusted measurements will improve when ED and ES phases are determined manually.

Limitations

First, a small patient population was examined; therefore data need to be validated by a larger patient group. Second, we didn’t compare our results to a gold standard. In many recent publications, MRI is considered the gold standard for obtaining LV functional parameters due to its excellent temporal resolution and superior tissue contrast between blood pool and myocardium. Many studies report a good agreement of cardiac MDCT compared to MRI [4, 5, 12, 13]. Therefore, we regarded the manual adjusted measurements as the reference standard in each software package. On the other hand, spatial resolution of CT datasets are superior compared to MRI, and several phantom studies show a better approximation of CT datasets to the physical volumes compared to MRI, therefore it is even questionable whether MRI should be the reference standard. Another limitation of our study is the number of frames per cardiac cycle; we used 10 phases per RR-interval, 20 phases could produce more accurate results. However, since we used the same dataset for each software tool, so this limitation may not affect the comparisons.

Summarizing the results of our study, it can be stated that although in many patients a small difference was observed and correlation between automatic and manually adjusted measurements was good, visual control in (semi-)automatic measurements is still mandatory because of a significant difference between at least two different LV functional
parameters in all software tools and several, clinically not acceptable outliers.

Acknowledgments The authors would like to thank Ronald Dob, Sigrid Brink and Jan Braaksma for their technical support and Stella Noach for reviewing this manuscript.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

1. White HD, Norris RM, Brown MA et al (1987) Left ventricular end-systolic volume as the major determinant of survival after recovery from myocardial infarction. Circulation 76(1):44–51
2. Hammermeister KE, DeRouen TA, Dodge HT (1979) Variables predictive of survival in patients with coronary disease: selection by univariate and multivariate analyses from the clinical, electrographic, exercise, arteriographic, and quantitative angiographic evaluations. Circulation 59:421–430
3. Shah PK, Maddahi J, Staniloff HM et al (1986) Variable spectrum and prognostic implications of left and right ventricular ejection fraction in patients with and without clinical heart failure after acute myocardial infarction. Am J Cardiol 58:387–393
4. Brodoefel H, Kramer U, Reimann A et al (2007) Dual-source CT with improved temporal resolution in assessment of left ventricular function: a pilot study. AJR 189:1064–1070
5. Yamamura M, Tadamura E, Kubo S et al (2005) Cardiac functional analysis with multi-detector row CT and segmental reconstruction algorithms: comparison with echocardiography, SPECT, and MR imaging. Radiology 234:381–390
6. Myerson SG, Bellenger NG, Pennell DJ (2002) Assessment of left ventricular mass by cardiovascular magnetic resonance. Hypertension 39(3):750–755
7. Sandstedt J, Lipke C, Beer M et al (2000) Age- and gender-specific differences in left and right ventricular cardiac function and mass determined by cine magnetic resonance imaging. Eur Radiol 10(3):438–442
8. Vogel-Clausen J, Finn JP, Gomes AS et al (2006) Left ventricular papillary muscle mass: relationship to left ventricular mass and volumes by magnetic resonance imaging. J Comput Assist Tomogr 30:426–432
9. Scheffel H, Alkadhi H, Plass A et al (2006) Accuracy of dual-source CT coronary angiography: first experience in a high pre-test probability population without heart rate control. Eur Radiol 16:2739–2747
10. Pugliese F et al (2006) Diagnostic accuracy of non-invasive 64-slice CT coronary angiography in patients with stable angina pectoris. Eur Radiol 16:575–582
11. Bastarrika G, Arraiza M, Pueyo JC et al (2008) Quantification of left ventricular function and mass in cardiac dual-source CT (DSCT) exams: comparison of manual and semiautomatic segmentation algorithms. Eur Radiol 18:939–946
12. Cury RC, Nieman K, Shapiro MD et al (2008) Comprehensive assessment of myocardial perfusion defects, regional wall motion, and left ventricular function by using 64-section multidetector CT. Radiology 248:466–475
13. Busch S, Johnson TR, Wintersperger BJ et al (2008) Quantitative assessment of left ventricular function with dual-source CT in comparison to cardiac magnetic resonance imaging: initial findings. Eur Radiol 18:570–575
14. Mahnken AH, Mühlenbruch G, Koos R et al (2006) Automated vs. manual assessment of left ventricular function in cardiac multidetector row computed tomography: comparison with magnetic resonance imaging. Eur Radiol 16:1416–1423
15. Boehm T, Alkadhi H, Roffi M et al (2004) Time-effectiveness, observer-dependence, and accuracy of measurements of left ventricular ejection fraction using 4-channel MDCT. Rofo Fortschr Geb Rontgenstr Neuen Bildgeb Verfahr 176:529–537
16. Juergens KU, Seifarth H, Range F et al (2008) Automated threshold-based 3D segmentation versus short-axis planimetry for assessment of global left ventricular function with dual-source MDCT. AJR 190:308–314
17. Bland JM, Altman DG (1986) Statistical methods for assessing agreement between two methods of clinical measurements. Lancet 1:307–310
18. Gerber TC, Kuko RS, Morin RL (2005) Techniques and parameters for estimating radiation exposure and dose in cardiac computed tomography. Int J Cardiovasc Imaging 21:165–176
19. Dulce MC, Mostbeck GH, Friese KK et al (1993) Quantification of the left ventricular volumes and function with cine MR imaging: comparison of geometric models with three-dimensional data. Radiology 188:371–376
20. Pluhans C, Keil S, Ocklenburg C et al (2009) Comparison of manual, semi- and fully automated heart segmentation for global left ventricular function in multidetector computed tomography. Invest Radiol 44(8):476–482
21. Mühlenbruch G, Das M, Hohl C et al (2006) Global left ventricular function in cardiac CT. Evaluation of an automated 3D region-growing segmentation algorithm. Eur Radiol 16:1117–1123
22. Krishnam MS, Tomassian A, Iv M et al (2008) Left ventricular ejection fraction using 64-slice CT coronary angiography and new evaluation software: initial experience. Br J Radiol 81(966):450–455
23. De Jonge GJ, Van Ooijen PMA, Van der Vleuten PA et al. (2009) The reliability of automatic measurement of left ventricular function with dual-source computed tomography datasets. Eur Radiol [epub ahead of print]