A study investigating variability of left ventricular ejection fraction using manual and automatic processing modes in a single setting

Authors: Philippa Bresser[a], Janeske de Beer[a,b] and Yolandi de Wet[a,c]

Department of Radiography, University of Pretoria

Running Title: Left ventricular ejection fraction inter-operator variability

Correspondence to:

P. Bresser
Department of Radiography
University of Pretoria
Steve Biko Academic Hospital
Level F5
Gezina
Pretoia
0002
Tel: +27(0)12-354-4348
pippa.bresser@up.ac.za

Grant/funding Support: None declared. Financial Disclosures: None declared.

Abstract

Purpose: A planar multi-gated cardiac blood pool acquisition is a non-invasive technique commonly used to measure left ventricular ejection fraction (LVEF). It is essential that the calculation of LVEF be accurate, repeatable and reproducible for serial monitoring of patients. Different processing modes may be used in calculating the LVEF which require various degrees of manipulation. In addition, different operators with varying levels of experience may process the same data set. It is not known whether the inter-operator variability of LVEF values within a single nuclear medicine department has the potential to affect the calculated LVEF and in turn affect patient management. The aim of the study was to determine variability of LVEF values among operators with different levels of experience using two processing modes.

Methods: A descriptive cross-sectional study was carried out in a single setting. Four operators with varying levels of experience analysed 120 left anterior oblique projections using manual and automatic processing modes to calculate the LVEF. Inter- and intra-operator correlation was determined.

Results: Manual processing showed moderate to strong agreement (r1=0.653) between operators. Automatic processing indicated almost perfect (r1=0.812) inter-operator correlation. Intra-operator correlation demonstrated a trend of decreasing variability between processing modes with increasing levels of experience.

Conclusion: Despite the overall inter-operator agreement, significant intra-operator variability was evident in results from operators with less experience. However, the discrepancies were such that the differences in LVEF would not

play a role in patient management. It is recommended that automatic processing be used for determining LVEF to limit inter-operator variability. Additionally operator experience should be considered in the absence of standardised processing protocols when different processing modes are available in a single setting.

Keywords: equilibrium radionuclide angiography, processing methods, intra-observer variability, left ventricular ejection fraction

[a] Department of Radiography, Faculty of Health Sciences, School of Healthcare Sciences, University of Pretoria, South Africa

[b] Dr G Lange Inc, Little Company of Mary Hospital, Pretoria, South Africa

[c] Department of Nuclear Medicine, Steve Biko Academic Hospital, Pretoria, South Africa

Correspondence to: Philippa Bresser, Department of Radiography, Faculty of Health Sciences, University of Pretoria, Private Bag x323, Arcadia, Pretoria, 0007, South Africa. Tel.: +27(0)12-354-4348. E-mail: pippa.bresser@up.ac.za

Introduction

A planar multi-gated cardiac blood pool acquisition (MUGA) is considered as an accurate, safe, non-invasive method commonly used to evaluate left ventricular ejection fraction (LVEF).[1-3] Any significant change in LVEF (decrease of 10%) from baseline is an early indicator of cardiac failure and may precede any symptoms of cardiac disease. [3-6] Early detection of changes allows for intensified monitoring to prevent further complications, initiation of preventative measures or may implicate changes in cancer treatments or patient management.[7,8] After gated image acquisition, regions of interest (ROIs) are drawn over the left ventricle at end systole and end diastole using processing software and algorithms are applied to calculate the LVEF.[9,10] Previously, only manual processing modes were used to determine LVEF.[1,11] Technological improvements led to the introduction of automatic and semi-automatic processing modes and the evolution of various processing software. However, these developments also increased the probability of variation in determining the LVEF. Sources of variability in manual and automatic processing modes include the following:

(i) Differences between software packages and applied algorithms to calculate LVEF;[1,2,4,12]

(ii) Human detection inaccuracies of the true end systole and end diastole images;[1,13]

(iii) Differences in software edge-detection algorithms;

(iv) Arrhythmias (e.g. ectopic beats or atrial fibrillation) which lead to artificially reduced counts in frames later in the cardiac cycle (end-diastole);[1]

(v) Inclusion of other anatomic structures within the left ventricle (LV) ROI due to poor positioning or anatomic variations; [1,8]

(vi) Poor labelling of the red blood cells leading to reduced count rates; [1]

(vii) Increased background counts leading to improper detection of the LV edge and reduced accuracy of the LVEF;[13]

(viii) Different operators with various levels of experience processing the same data set which formed the basis of the current study.[1-3,14]

The extent of variability in LVEF results in a particular setting should be considered where multiple operators and processing modes are available.[13,15] The calculation of LVEF must be accurate, repeatable and reproducible, as serial monitoring of LVEF variations has the potential to affect patient treatment and management.[4,11] 'Repeatability' of LVEF is an expression of the variability of repeated measurements of the same acquisition as opposed to 'reproducibility' which includes both repeated acquisition and measurement preferably at different times and places.[14] Inter- and intra-operator variability in MUGA processing and data manipulation is a principal concern in nuclear medicine departments where the possibility of errors or discrepancies should be kept to a minimum.[2-4]. Hains et al compared the LVEF values from three operators using three different processing modes (manual, semi-automated and regional) and established that there was no significant inter-operator variability.[2] This was supported by Bailey and Bailey in a study across multiple settings. [3] However, these studies did not consider the different levels of experience of the operators in determining inter-operator variability. The intra-operator variability

determined by Bailey and Bailey did include the different levels of experience of the operators but the results were based on the use of a single processing mode. It is thus contended that inter- and intra-operator differences in LVEF results may occur where operators in the nuclear medicine department have varying levels of clinical experience in processing MUGA examinations using different processing modes. The purpose of this study was to investigate the extent of inter- and intra-operator variability in calculating LVEF through automatic and manual processing modes with multiple operators within a single setting.

Methods

A descriptive cross-sectional study was carried out in a nuclear medicine department in Pretoria, South Africa where retrospective analysis of MUGA examinations was undertaken. Ethical approval to conduct this study was granted by the Ethics Committee of the Faculty of Health Sciences, at the University of Pretoria. The data collected from the picture-archiving system of the hospital consisted of the LAO projection of all patients that underwent a MUGA examination from October to April 2013 until a consecutive sample of 120 data sets was obtained. The images had been acquired using the General Electric (GE), Millennium, single head gamma camera (GE Healthcare) fitted with a low-energy, high-resolution collimator. All studies were acquired with a 64x64 matrix in the LAO projection using 24 frames.

Four operators with varying levels of experience and clinical expertise in the processing of nuclear medicine examinations volunteered to process the data. The 4 operators included one nuclear medicine registrar (doctor) with four years' experience in reporting nuclear medicine examinations, one nuclear medicine radiographer with 12 years' experience post-graduation (senior), one junior nuclear medicine radiographer with 3 years' experience post-graduation and one nuclear medicine student in the first year of training. The 4 operators were blinded to the patient information and were blinded to the results of the other operators. The data was processed using Xeleris 2.0 software (GE Healthcare). Each operator processed 120 data sets both manually and automatically to determine the LVEF. Manual processing required the operator to visually identify and manually draw the region of interest around the edge of the left ventricle. Automatic processing required the operator to select the centre of the left ventricle and the computer identified the edge of the ventricle through the use of an edge-detection algorithm. Manual and automatic processing was performed at different time points in order to minimise the risk of operator bias in trying to replicate results using different processing modes. The results of the manual and automatic LVEF values were recorded on separate data collection sheets. Upon completion of processing using one mode, the data collection sheet was placed in a sealed box whereafter processing using the other mode was performed. The authors were not involved in the processing or collecting of data. Procedural bias was avoided in that operators could select the two time points that was most appropriate for them to complete the processing depending on their work load and personal

preference. Operators were volunteers and were not given any incentives to participate.

Inter-operator variability was considered as the amount of variation between LVEF results obtained by the different operators processing the same data set. Intra-operator variation was considered as the variation of LVEF value calculated by one operator processing the same data set using two different processing modes.

Inter-operator correlation coefficient (r1) was calculated for manual and automatic processing modes to determine the agreement of LVEF values among the 4 operators. The variability between manual and automatic LVEF among the operators was described using summary statistics. Additionally, intra-operator correlation was determined using the Wilcoxon signed rank test to determine any discrepancies when the same data was processed multiple times by an individual, using different processing modes.

Results

Inter-operator intra-class correlation coefficient (r1) was calculated with a 95% confidence interval (CI) as seen in Table 1. Manual processing showed a significant moderate to strong agreement between operators. Automatic processing indicated an almost perfect and significant inter-operator correlation. LVEF values ranged from 20% to 84%. Table 2 illustrates the variability of LVEF values among the operators. Intra-operator correlation among the four operators was determined using the Wilcoxon signed rank test. The LVEF values from the doctor and senior radiographer did not show a significant

difference ($p>0.05$) between the manual and automatic results. A significant intra-operator difference existed, however, between the manual and automatic LVEF values obtained by the student and junior radiographer. There was a 6.97% deviation of LVEF between the manual and automatic processing modes for the student and 1.38% for the junior radiographer (Fig 1).

Discussion

The choice of processing mode to quantify the LVEF is operator dependent and based on personal preference and/or departmental protocol.[1,2] ROIs can be drawn manually, semi-automatically or automatically.[1,2,12] Automatic ROIs require minor manipulation from the operator. However, manually drawn ROIs require major manipulation, which can lead to increased possibilities of error that are amplified when different operators with various levels of experience process the same data.[1-3,5,14] The results from the current study indicate that there was overall inter-operator agreement between manual and automatic processing modes which is similar to other studies. [2,3,16] However, automatic processing yielded almost perfect agreement of LVEF values between operators when compared to manual processing. This is similar to the results from Hains et al. who established that although there was no significant inter-operator variability, the automated processing mode gave more consistent results.[2] Bailey and Bailey also reported that manually drawn regions, increase random error and decrease accuracy and reproducibility. These authors found that the inter-operator variability was within range with a mean deviation of 3,4% upon determination of LVEF reproducibility across multiple departments where

each department used one processing mode.[3] The mean inter-operator deviation between manual and automatic processing for the current study was 1.33%. This reduction could be explained by the difference in the respective study designs. The current study evaluated inter- and intra-operator discrepancies within a single department as opposed to multiple departments. Intra-operator correlation of manual and automatic LVEF values showed a significantly large intra-operator difference existed between the student and junior results. There was no significant intra-operator difference of LVEF values calculated by the senior and doctor. Hence, there was a tendency of less intra-operator variability between manual and automatic LVEF values with increasing levels of operator experience. Although the intra-operator variability did not exceed 10%, the difference between the manual and automatic processing results is of concern among operators with less experience. Hole et al also reported greater variation in results when comparing intra-operator repeatability across various hospitals and attributed it to an operator with the least experience.[14] The intra-operator variability analysed across 4 departments using results of 3-7 operators in the study by Bailey and Bailey had a much wider range (from 0,1 – 27,9%) than the current study. This could be attributed to the slightly increased number of operators included per department as well as the variation in years' experience of the operators at each department.. In contrast to the results from the current study, Bailey and Bailey did not find a correlation between operator experience and average deviation in LVEF after multiple processing attempts. This discrepancy can be attributed to the fact that variability was measured after operators performed the processing using one

processing mode whereas in the current study the deviation was determined by comparing two processing modes. This study highlights the importance of informing the nuclear medicine departments of discrepancies that exist in determining LVEF especially in cases where different processing modes are used. More particularly where more than one operator may process the images, the operator experience should be considered in the absence of standardised processing protocols. Although a limitation of this study was the use of one operator in each category resulting in inclusive bias, the aim of the study was to determine the extent of inter- and intra-operator variability within a single setting. Hence, the results may not be generalizable to nuclear medicine departments that have larger numbers of operators or operators with similar years' experience.

Conclusion

Varying levels of operator experience in processing techniques and clinical practice can lead to discrepancies in the processing technique of MUGA examinations. This study aimed to investigate the extent of inter- and intra-operator variability in calculating LVEF through automatic and manual processing modes with multiple operators within a single setting. It was found that automatic processing should be used for determining LVEF to limit inter-operator variability and thus reduce the risk of affecting the serial monitoring and management of patients. Despite the overall inter-operator agreement, there was a trend of less intra-operator variability between the two processing modes with increasing years of experience. Hence operator experience should

be considered when different processing modes are available. From a quality perspective, the monitoring of potential discrepancies that exist between operators or within operators who use different processing modes is of importance to improve repeatability of results, thus minimising the potential impact on diagnosis and patient management.

Conflict of interest statement

The authors have no conflicts of interest.

Acknowledgements

To Barbara English of the University of Pretoria's Research Office (Faculty of Health Sciences) for her editing of this article. To Dr CM Makanjee and Ms M Dutton for their input and insight.

References

1. Foley T, Mankad S, Anavekar N, Bonnichsen C, Morris M, Miller T, et al. Measuring left ventricular ejection fraction - techniques and potential pitfalls. *Eur Cardiol* 2012;**8**(2):108.
2. Hains A, Imad A, Hinge D, Lahiri A, Raftery E. Radionuclide left ventricular ejection fraction: a comparison of three methods. *Br Heart J* 1987;**57**:242.
3. Bailey E, Bailey D. Results from an Australian and New Zealand audit of left ventricular ejection fraction from gated heart pool scan analysis. *Nucl Med Commun*. 2012;**33**(1):102.
4. Davis BA, O'Sullivan C, Jarritt PH, Porter JB. Value of sequential monitoring of left ventricular ejection fraction in the management of thalassemia major. *Blood* 2004 July 01;**104**(1):263-269.
5. Christian P, Waterstram-Rich K, editors. Nuclear medicine and PET/CT technology and techniques. 7th ed. St. Louis. Missouri: Mosby Elsevier; 2012. p.529
6. de Geus-Oei L, Mavinkurve-Groothuis AMC, Bellersen L, Gotthardt M, Oyen WJG, Kapusta L, et al. Scintigraphic Techniques for Early Detection of Cancer Treatment–Induced Cardiotoxicity. *J Nucl Med*. 2011;**52**(4):560-571.
7. Altena R, Perik P, van Veldhuisen D, de Vries E, Gietema J. Cardiovascular toxicity caused by cancer treatment: strategies for early detection. *Lancet Oncol*. 2009;**10**:391-99.
8. Slart R, Poot L, Piers D, van Veldhuisen D, Nichols K, Jager P. Gated blood-pool SPECT automated versus manual left ventricular function calculations. *Nucl Med Commun.* 2004;**25**(1):75.

9. Abe M, Kazatani Y, Fukuda H, Tatsuno H, Habara H, Shinbata H. Left ventricular volumes, ejection fraction, and regional wall motion calculated with gated technetium-99m tetrofosmin SPECT in reperfused acute myocardial infarction at super-acute phase: comparison with left ventriculography. *J Nucl Cardiol.* 2000;**7**(6):569-574.
10. Scheiner J, Sinusas A, Wittry M, Royal H, Machac J, Balon HL,O. Society of Nuclear Medicine procedure guidelines manual. Gated equilibrium radionuclide venticulography. Version 3. 2002; Available at: http://www.snmmi.org/ClinicalPractice/content.aspx?ItemNumber=6414. Accessed January 2014.
11. Lum D, Coel M. Comparison of automatic quantification software for the measurement of ventricular volume and ejection fraction in gated myocardial perfusion SPECT. *Nucl Med Commun.* 2003;**24**(3):259-266.
12. Lin X, Xu H, Zhao X, Folks RD, Garcia EV, Soman P, et al. Repeatability of left ventricular dyssynchrony and function parameters in serial gated myocardial perfusion SPECT studies. *J Nucl Cardiol.* 2010;**17**(5):811-816.
13. Hiscock S, Evans M, Morton R, Hall D. Investigation of normal ranges for left ventricular ejection fraction in cardiac gated blood pool imaging studies using different processing workstations. *Nucl Med Commun.* 2008;**29**(2):103.
14. Hole T, Åsberg A, Graven T, Lied A, Morstøl TH, Skjaerpe T. Intra- and interhospital repeatability of radionuclide left ventricular ejection fraction in post-infarction patients. *Scand Cardiovasc J.* 2001;**35**(1):35-39.

15. Skrypniuk JV, Bailey D, Cosgriff PS, Fleming JS, Houston AS, Jarritt PH, et al. UK audit of left ventricular ejection fraction estimation from equilibrium ECG gated blood pool images. *Nucl Med Commun.* 2005;**26**(3):205.
16. Kozelka J, Voslar A. Manual-processed vs. automated-processed left ventricular ejection fractions in MUGA studies. *J Nucl Med.* 2013;**54**(Suppl 2):2735.

Tables

Table 1

Inter-operator correlation of LVEF for manual and automatic processing

	ICC	95% CI ranges		p-value
Manual	r1 = 0.65	0.49	0.77	0.00*
Automatic	r1 = 0.81	0.76	0.86	0.00*

*p≤0,05 considered significant

Table 2

Intra-operator variability of manual and automatic LVEF values

	Student	Junior	Senior	Doctor
Manual LVEF (%)				
Min	20	32	34	32
Max	76	74	77	76
Mean	53.48	58.87	60.30	59.98
Standard deviation	10.88	8.31	7.97	8.21
Automatic LVEF (%)				
Min	39	37	32	40
Max	84	78	78	77
Mean	60.45	60.25	59.47	59.99
Standard deviation	8.19	8.18	9.95	7.87
Intra-operator deviation (%)	6.97	1.38	-0.83	0.01

Figure Legends

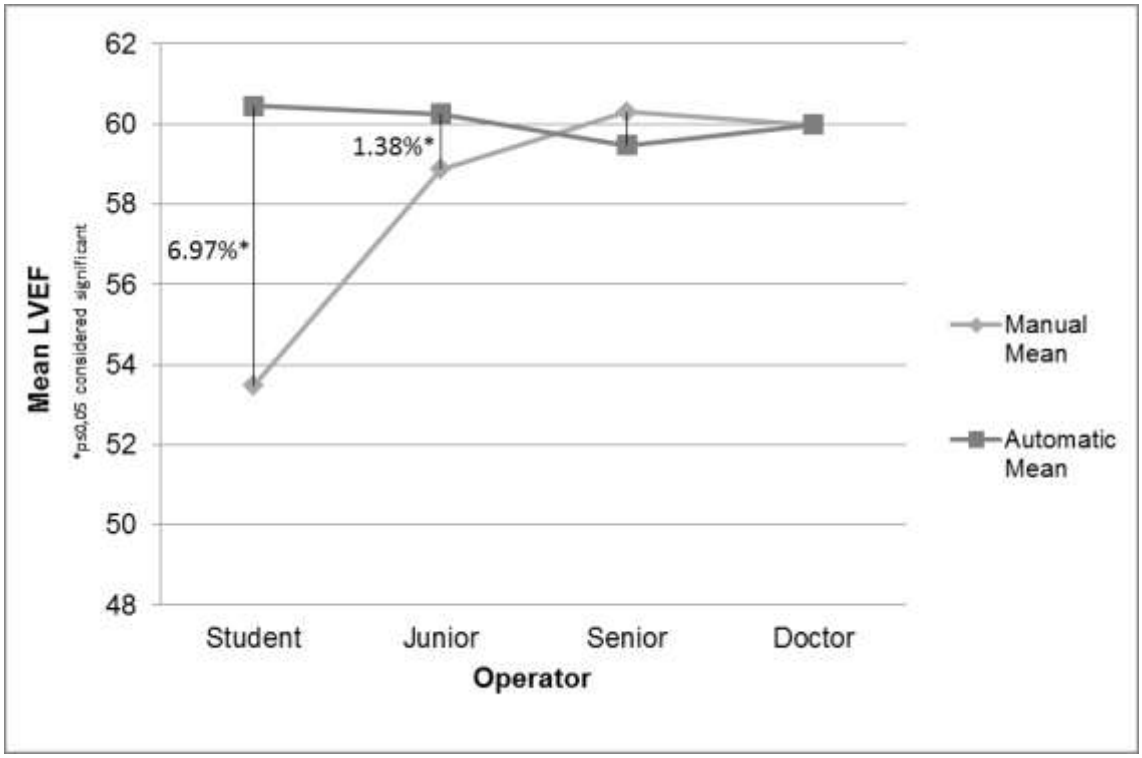

Figure 1. Inter-operator mean LVEF and intra-operator deviation of manual and automatic processing for the four operators