ORIGINAL ARTICLE

Intersoftware variability impacts classification of cardiac PET exams

Joaquim Barreto Oliveira,[a] Yew Min Sen, MD,[b] and Kshama Wechalekar, MD[c]

[a] Faculty of Medical Sciences, State University of Campinas (UNICAMP), Campinas, SP, Brazil
[b] Tan Tock Seng Hospital, Singapore, Singapore
[c] Nuclear Medicine Department, Royal Brompton & Harefield Trust, London, UK

Received Oct 18, 2017; accepted Sep 2, 2018
doi:10.1007/s12350-018-1444-z

Background. **Myocardial perfusion imaging (MPI) with ^{82}Rb PET/CT is increasingly utilized in the evaluation of coronary artery disease with high diagnostic accuracy. Various softwares for data processing have been developed over the years with conflicting data regarding their reproducibility. In this study, we compared the quantitative results of myocardial perfusion and exam classification from three different softwares.**

Methods. **Data from consecutive patients who have undergone rest/stress ^{82}Rb PET/CT MPI at the Royal Brompton & Harefield Trust, London, were analyzed. All data were processed using the Corridor 4DM (Invia, Ann Arbor, Michigan, USA), QPET (Cedars-Sinai, Los Angeles, California, USA), and SyngoMBF (Siemens Healthineers, Erlangen, Germany). The software packages addressed Lortie tracer kinetic model and region of interest (ROI) extraction correction option.**

Statistics. **A repeated-measures ANOVA with a Greenhouse–Geisser correction was performed with post hoc tests using Bonferroni correction. For intersoftware variability, Pearson correlation and intraclass correlation coefficients (ICC) were calculated. Bland–Altman assessed limit of agreement. Cohen's Kappa assessed agreement in the classification of exams as normal or abnormal using an MFR cut-off value of 2.0. A *P* value of less than 0.05 was considered statistically significant.**

Results. **Data from 55 patients were analyzed. The mean values of myocardial blood flow (MBF) and myocardial perfusion reserve (MFR) were statistically significantly different among the softwares ($P < 0.05$). Corridor4DM had considerably lower values of MFR and classified a more substantial number of exams as abnormal (MFR: 2.21 ± 0.7, 2.4 ± 0.8, and 1.98 ± 0.8; and 18, 15, and 31 exams were abnormal for Syngo, QPET, and Corridor4DM, respectively). Accordingly, kappa agreement was moderate for Syngo vs QPET ($k > 0.5$), but minimal for Corridor4DM in comparison to its pairs ($k < 0.4$).**

Conclusion. **Users should be cautious when using different software interchangeably as systematic differences amongst them may introduce more extensive quantitative variation which could be clinically significant. (J Nucl Cardiol 2019;26:2007–12.)**

Key Words: Nuclear cardiology • MPI • risk stratification

See related editorial, pp. 2013–2017

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s12350-018-1444-z) contains supplementary material, which is available to authorized users.

The authors of this article have provided a PowerPoint file, available for download at SpringerLink, which summarises the contents of the paper and is free for re-use at meetings and presentations. Search for the article DOI on SpringerLink.com.

Reprint requests: Joaquim Barreto Oliveira, Faculty of Medical Sciences, State University of Campinas (UNICAMP), Rua Tessália Vieira Camargo 126, Cidade Universitária, Campinas, SP, Brazil; *joaquimbarretoantunes@gmail.com*

1071-3581/$34.00

INTRODUCTION

Myocardial perfusion imaging (MPI) with ^{82}Rb PET/CT is increasingly utilized in the evaluation of patients with suspected or known coronary artery disease (CAD). Importantly, its diagnostic accuracy is comparable to coronary angiography and higher than single photon emission computed tomography (SPECT) or stress echocardiography.[1] Apart from providing attenuation corrected perfusion and gated values, it also provides a quantitative assessment of myocardial blood flow (MBF), which has added prognostic value over perfusion data alone.[2]

Concerning its clinical importance, various quantitative softwares for data processing have been developed over the years.[3] They differ in terms of algorithms for sampling counts in the myocardium and blood pool. There are conflicting data with regard to the reproducibility between various software packages regarding myocardial perfusion, with some demonstrating good agreement and others showing differing results.[4] However, these studies were performed using earlier software versions and may not reflect recent refinements in software algorithms, use of newer stress agents such as regadenoson and improvements in PET MPI technique.

In this study, we aim to compare the agreement between the latest version of Corridor4DM v2017 (Invia, Ann Arbor, Michigan, USA), Quantitative PET (QPET, Cedars-Sinai, California Los Angeles, USA) and Syngo MBF (Siemens Healthineers, Erlangen, Germany). We evaluated the quantitative myocardial perfusion in patients undergoing rest/stress PET/CT, with either adenosine or regadenoson as stress agent.

MATERIALS AND METHODS

Inclusion Criteria and PET Protocol

We retrospectively analyzed data from consecutive patients who have undergone rest/stress ^{82}Rb PET/CT MPI at the Royal Brompton & Harefield Trust, London, between August and December 2016. These patients had to be fasted for a minimum of 2 hours and abstain from caffeine for at least 24 hours before the test. A whole-body PET/CT scanner (Biograph mCT, Siemens Healthcare) was used for image acquisition. Attenuation correction CT scan was performed before rest and stress imaging. All patients underwent pharmacological stress with either slow intravenous bolus of regadenoson 400mcg or infusion adenosine 140 μg kg^{-1} $minutes^{-1}$ for 6 minutes. Emission images were obtained after intravenous administration of 1100 MBq of Rb-82 at both rest and stress, starting at 120 seconds after completion of Rb-82 infusion and continued for 7 minutes.

Data Processing

All cardiac PET/CT data were processed using Corridor4DM v2017 (Invia, Ann Arbor, Michigan, USA), Quantitative PET (QPET, Cedars-Sinai, California Los Angeles, USA), and Syngo Dynamic PET MBF (Syngo MBF, Siemens Healthineers, Erlangen, Germany). Stress and rest MBF values in mL $minutes^{-1}$ g^{-1} and myocardial flow reserve (MFR) were determined for all subjects. Data were processed using the respective software by a single observer blinded to clinical information. End-user intervention such as manual adjustment of myocardial contours was avoided unless there was failure of automated myocardial edge detection that had resulted in erroneous measurements.

For each dataset, the arterial input function was quantified from dynamic images using the semi-automatically generated 3-dimension region of interest (ROI). Later, the data were fit to a 1-tissue compartment to estimate k1 (ml $minutes^{-1}$ g^{-1}) and k2 ($minutes^{-1}$), which denote the extraction and egress rates of transport between the metabolically trapped space (myocardial tissue) and the freely diffusible space (ventricular blood pools), respectively.[5] Absolute MBF was quantified using the kinetic rates according to Lortie et al. validated extraction model.[6]

Noteworthy, for all software programs, myocardial segmentation was performed automatically with differing methods for time-activity curve achievement. Briefly, for Corridor 4DM time activity curve was obtained from an LV input function derived from a ROI centered between the left ventricle and left atrium, and MBF is computed for each of the 17 segments of the polar map. Differently, in QPET a cylindrical ROI was automatically placed in the middle of the valve plane, with a 1 × 2 cm length oriented along the long axis of the heart. A total of 70 myocardial regions are considered and results are presented by averaging MBF in 17-segment polar map. Finally, for Syngo MBF cylindrical ROI was placed in the middle of the LV in the basal region, with 1.5 × 2 cm length. MBF is computed on 505 segments and reported by averaging for each 17-segment polar map segments.

Rate Pressure Product Correction

Rate pressure product (RPP) correction was applied for rest MBF values, and the corrected global MFR results were noted. RPP correction was generated automatically in Corridor 4DM and QPET once the patient's resting heart rate and blood pressure were keyed into the programs. Manual calculation by formula (resting MBF divided by RPP and multiplied by a linear factor of 8000) was required in Syngo MBF.

Statistical Analysis

Statistical analysis was performed using SPSS, version 20.0 (IBM, Chicago, IL, USA). Continuous variables are presented as mean ± SD. Means were compared by repeated-measures ANOVA with a Greenhouse–Geisser correction with post hoc tests using Bonferroni correction. For characterization of intersoftware variability, Pearson correlation and intraclass correlation coefficients (ICC) were calculated. Bland–Altman assessed limit of agreement as the mean ± coefficient of repeatability (RPC = 1.96 x SD). Chi-square test was used when comparing categorical data. Mean differences were compared to 0 by one-sample *t* test. Finally, agreement in the classification of all patients using a MFR cut-off value of 2.0 as the threshold between normal and abnormal values was assessed by Cohen's Kappa agreement. For all analysis, a *P*-value of less than 0.05 was considered statistically significant.

RESULTS

Data from 55 patients were analyzed. Population characteristics are summarized in Table 1. The mean values obtained by the software packages were significantly different (Table 2). Pairwise analysis showed significantly lower values of stress and reserve flows for Corridor4DM in comparison to its pairs. Syngo and QPET generated comparable results. These differences were consistent across vascular territories (Table 2).

All software packages showed moderate to strongly correlated values, as determined by Pearson correlation coefficient (r) and intraclass correlation coefficient (ICC) (Table 3). Bland–Altman analysis showed biases of no more than 0.4, 0.37, and 0.09 for MFR, stress, and rest MBF, respectively (Table 4). Moreover, Bland–Altman showed that values obtained by Corridor4DM were consistently lower than its pairs, as demonstrated by predominantly positive differences (Figure 1).

As a consequence, for an MFR cut-off value of 2 as the threshold between normal and abnormal values, Corridor4DM classified a larger number of exams as abnormal (15, 18, and 31 for Syngo, QPET, and Corridor4DM, respectively). Accordingly, kappa agreement was higher between Syngo and QPET (53%) than it was when Corridor4DM was compared to its pairs (34-38%) (Table 5).

Table 1. Population characteristics

Characteristics	
Age, years	65.4 ± 10
Male, %	38 (69)
Current smoking, %	5 (9)
Diabetes, %	20 (36)
Dyslipidemia, %	33 (60)
Hypertension, %	30 (56)
Obesity, %	21 (42)
BMI, kg m^{-2}	28.6 ± 8
Heart rate, bpm	64 ± 14
SBP, mmHg	133 ± 21
Rate pressure product, bpm mmHg	8760 ± 3249

Numbers presents are mean ± standard deviation (SD) for continuous variable and number (%) for binary and categorical variables
BMI, body mass index; *bpm*, beats per minute; *SBP*, systolic blood pressure

DISCUSSION

Our study compared data from 55 consecutive patients submitted to ^{82}Rb PET/CT analyzed by three software programs. All packages used the same kinetic model and analysis was made by a single observer blinded to clinical information. Our main findings are as follows: (i) Corridor4DM had consistently lower values of MFR and classified a larger number of exams as abnormal; (ii) SyngoMBF and QPET results were overall comparable.

Our study found significant differences between Corridor4DM and its pairs. This software generated consistently lower values of flow, hence classifying a larger number of exams as abnormal when compared to QPET and SyngoMBF. Our findings are consistent with previous study by Tahari et al.[5], who firstly demonstrated significant differences in stress MBF and overall exam classification when comparing Corridor4DM to Flow Quant and Munich Heart softwares. Furthermore, it may be emphasized that a significant decrease in flow quantification was observed when ROI method was applied to Corridor4DM, what is consistent with our findings that C4DM-ROI presented lower values of flow when compared to its pairs.[5]

Many are the sources of variation that could explain the reported results. Of note, besides kinetic model intersoftware variability also relies on each package automated quantitative methods of finding and segmenting the myocardial and blood pool.[7] In this, while Corridor4DM places ROI between LV and left atrium, QPET and SyngoMBF place ROI in the middle of LV and in the middle of the mitral valve plane, respectively. Besides, while Corridor4DM calculates MBF considering the 17 segments of the polar map, QPET and Syngo consider the average of over 70 and 505 myocardial segments. Moreover, variation may be introduced by manual intervention, which is often required when strong subdiaphragmatic uptake weakens automation accuracy, and may widen intersoftware variation.[7]

Finally, the agreement between QPET and SyngoMBF found in our study is in accordance to previous results reported by deKemp et al.[6] For both studies, stress and reserve flows were comparable between softwares for both global and regional flows, when using Lortie kinetic modeling and ROI.[6] In fact, while deKemp et al.[6] reported differences of no more than 0.28 ml $minutes^{-1}$ g^{-1}, the mean difference between

Table 2. Mean values of flow

	Syngo	QPET	C4DM	ANOVA[a]	S × Q[b]	S × C[b]	Q × C[b]
Global							
Rest	1.1 ± 0.3	1 ± 0.2	1.1 ± 0.3	0.015	0.15	0.809	0.186
Stress	2.4 ± 0.9	2.3 ± 0.8	2.1 ± 0.7	0.001	0.204	0.001	0.001
MFR	2.21 ± 0.7	2.4 ± 0.8	1.98 ± 0.8	0.001	0.35	0.018	0.001
LAD							
Rest	1.08 ± 0.28	1.04 ± 0.25	1.19 ± 0.31	0.001	0.03	0.01	0.001
Stress	2.42 ± 0.87	2.32 ± 0.77	2.06 ± 0.71	0.001	0.19	0.001	0.001
MFR	2.24 ± 0.74	2.33 ± 0.78	1.84 ± 0.74	0.001	0.66	0.001	0.001
LCx							
Rest	1.08 ± 0.28	0.99 ± 0.26	1.16 ± 0.39	0.004	0.02	0.204	0.004
Stress	2.31 ± 0.86	2.22 ± 0.79	2.01 ± 0.76	0.001	0.24	0.001	0.001
MFR	2.17 ± 0.74	2.32 ± 0.93	1.87 ± 0.74	0.001	0.19	0.001	0.001
RCA							
Rest	1.06 ± 03	0.97 ± 0.28	1.10 ± 0.29	0.004	0.03	0.29	0.02
Stress	2.57 ± 1.05	2.45 ± 0.88	2.43 ± 1.04	0.085	0.15	0.17	0.10
MFR	2.49 ± 0.98	2.68 ± 1.01	2.23 ± 0.94	0.001	0.22	0.004	0.001

[a]One-Way ANOVA for Repeated Measures with Greenhouse-Geisser correction and Bonferroni post hoc
[b]Pairwise comparison with Bonferroni adjustment, *P* values (*S*, SyngoMBF; *Q*, QPET; *C*, Corridor 4DM)

Table 3. Correlation between values

	Syngo vs QPET		QPET vs Corridor		Syngo vs Corridor	
	r[a]	ICC[b]	*r*[a]	ICC[b]	*r*[a]	ICC[b]
Global						
MFR	73	82 (69-90)	68	78 (62-88)	69	75 (40-88)
Stress	91	95 (91-97)	88	88 (49-95)	87	89 (67-96)
Rest	59	72 (50-80)	84	91 (85-95)	46	62 (36-78)
LAD						
MFR	74	85 (75-91)	73	78 (40-90)	61	67 (23-84)
Stress	89	94 (90-97)	78	82 (54-91)	58	85 (68-92)
Rest	63	75 (55-85)	85	91 (82-95)	49	59 (25-77)
LCx						
MFR	78	86 (75-92)	92	92 (50-97)	72	76 (40-89)
Stress	90	95 (91-97)	93	93 (60-97)	89	93 (81-97)
Rest	64	76 (56-86)	66	76 (59-86)	46	55 (28-74)
RCA						
MFR	73	84 (56-82)	83	89 (78-94)	69	77 (51-88)
Stress	91	94 (89-96)	86	92 (87-96)	87	92 (86-95)
Rest	57	70 (49-83)	83	90 (84-94)	51	63 (34-79)

[a]Pearson correlation coefficient (%)
[b]Intraclass correlation coefficient (95%CI)
For ICC, $P < 0.01$ for all values

the referred programs was of 0.09 ml minutes^{-1} g^{-1} in our study.

In conclusion, in the absence of a comparable gold-standard, this study is not designed to determine which method is the most accurate. However, our results support that cut-off values for perfusion imaging should be adequately used according to the chosen method as previously proposed and that these intersoftware

Table 4. Bland-Altman results

	MFR			sMBF			rMBF		
	S vs Q	S vs C	Q vs C	S vs Q	S vs C	Q vs C	S vs Q	S vs C	Q vs C
Bias	0.19	− 0.22	− 0.42	− 0.09	− 0.37	− 0.27	− 0.09	− 0.02	0.07
SD	0.55	0.58	0.62	0.36	0.41	0.38	0.23	0.15	0.27
High LOA	1.27	0.91	0.80	0.62	0.43	0.47	0.36	0.27	0.60
Low LOA	− 0.89	− 1.36	− 1.64	− 0.80	− 1.17	− 1.01	− 0.54	− 0.31	− 0.46

Bias, mean difference, *SD*, difference standard deviation, *LOA*, limits of agreement as the mean difference ± coefficient of repeatability (RPC = 1.96 × SD)

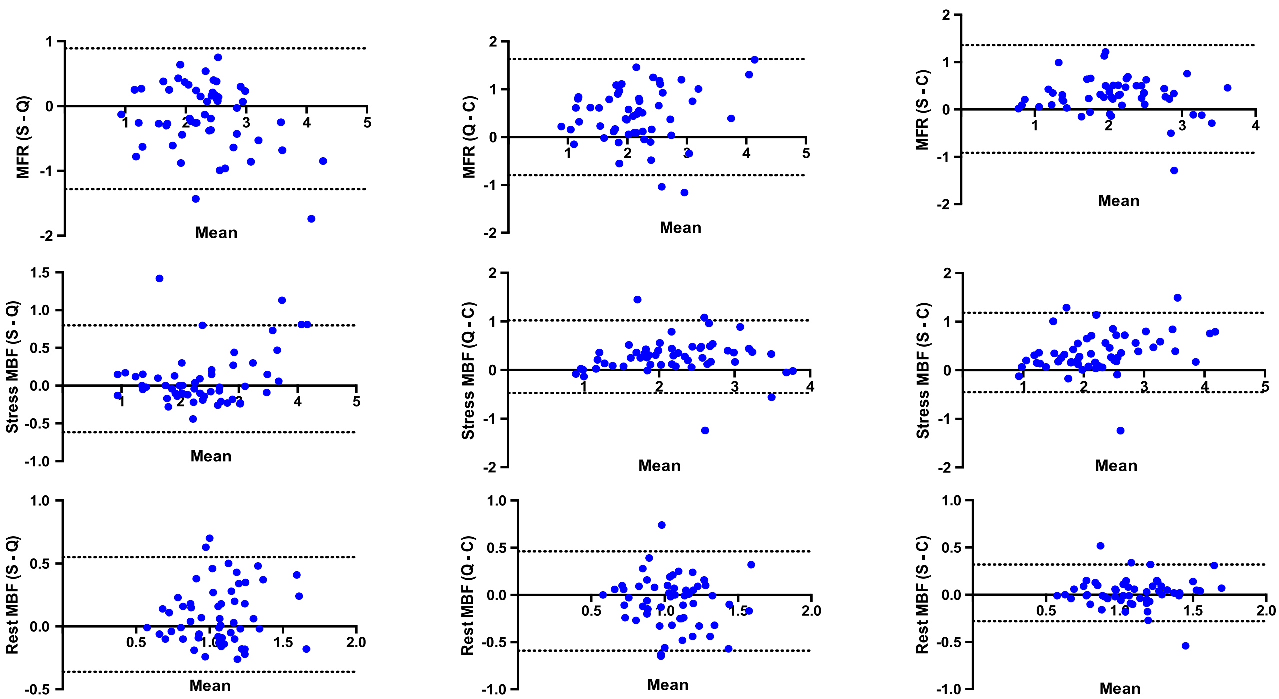

Figure 1. Bland–Altman plots illustrating differences in calculated reserve, stress, and rest flows (upper to lower panel) for different software combinations (left to right: S × Q, Q × C and S × C). Solid black lines are the mean differences, and dotted black lines are limits of agreement, calculated as mean difference ± 1.96SD. *S*, Syngo MBF; *Q*, QPET; *C*, Corridor4DM.

Table 5. Agreement of exam classification

	Exam classification		Agreement		
	Normal[a]	Abnormal[a]	S × Q[b]	S × C[b]	Q × C[b]
Syngo	37	18	**53%**	**34%**	**38%**
QPET	40	15			
C4DM	24	31			

[a]number of exams classified as normal (MFR ≧ 2) or abnormal (MFR < 2) by each program
[b]Cohen's Kappa agreement for exam classification among softwares
C4DM, Corridor4DM; *S*, Syngo MBF; *Q*, QPET; *C*, Corridor 4DM

differences should be kept in mind when comparisons between different studies are made.[8]

NEW KNOWLEDGE GAINED

Quantitative variation introduced by different software impacts exam classification.

CONCLUSION

Users should be cautious when using different software interchangeably as systematic differences amongst them may introduce wider quantitative variation which could be clinically significant.

Disclosure

No financial disclosures.

References

1. Sampson UK, Dorbala S, Limaye A, Kwong R, Di Carli MF. Diagnostic accuracy of rubidium-82 myocardial perfusion imaging with hybrid positron emission tomography/computed tomography in the detection of coronary artery disease. J Am Coll Cardiol. 2007;49:1052-8.
2. Farhad H, Dunet V, Bachelard K, Allenbach G, Kaufmann PA, Prior JO. Added prognostic value of myocardial blood flow quantitation in rubidium-82 positron emission tomography imaging. Eur Heart J Cardiovasc Imaging. 2013;14:1203-10.
3. Ziadi MC, Dekemp RA, Williams KA, et al. Impaired myocardial flow reserve on rubidium-82 positron emission tomography imaging predicts adverse outcomes in patients assessed for myocardial ischemia. J Am Coll Cardiol. 2011;58:740-8.
4. Bravo PE, Chien D, Javadi M, Merrill J, Bengel FM. Reference ranges for LVEF and LV volumes from electrocardiographically gated ^{82}Rb cardiac PET/CT using commercially available software. J Nucl Med. 2010;51:898-905.
5. Tahari AK, Lee A, Rajaram M, et al. Absolute myocardial flow quantification with (82)Rb PET/CT: Comparison of different software packages and methods. Eur J Nucl Med Mol Imaging. 2014;41:126-35.
6. Dekemp RA, Declerck J, Klein R, et al. Multisoftware reproducibility study of stress and rest myocardial blood flow assessed with 3D dynamic PET/CT and a 1-tissue-compartment model of ^{82}Rb kinetics. J Nucl Med. 2013;54:571-7.
7. Nesterov SV, Deshayes E, Sciagra R, et al. Quantification of myocardial blood flow in absolute terms using (82)Rb PET imaging: The RUBY-10 Study. JACC Cardiovasc Imaging. 2014;7:1119-27.
8. Herzog BA, Husmann L, Valenta I, et al. Long-term prognostic value of 13 N-ammonia myocardial perfusion positron emission tomography added value of coronary flow reserve. J Am Coll Cardiol. 2009;54:150-6.