Evaluation of the effect of myocardial segmentation errors on myocardial blood flow estimates from DCE-MRI

J Biglands[1], D Magee[2], R Boyle[2], A Larghat[3], S Plein[3] and A Radjenović[4]

1 Division of Medical Physics, University of Leeds, UK
2 School of Computing, University of Leeds, UK
3 Division of Cardiovascular and Neuronal Remodeling, University of Leeds, UK
4 Section of Musculoskeletal Disease, NIHR Leeds Musculoskeletal Biomedical Research Unit, UK

E-mail: jdb@medphysics.leeds.ac.uk

Abstract. Quantitative analysis of cardiac dynamic contrast enhanced magnetic resonance imaging (DCE-MRI) perfusion datasets is dependent on the drawing (manually or automatically) of myocardial contours. The required accuracy of these contours for myocardial blood flow (MBF) estimation is not well understood. This study investigates the relationship between myocardial contour errors and MBF errors. Myocardial contours were manually drawn on DCE-MRI perfusion datasets of healthy volunteers imaged in systole. Systematic and random contour errors were simulated using spline curves and the resulting errors in MBF were calculated. The degree of contour error was also evaluated by two recognized segmentation metrics. We derived contour error tolerances in terms of the maximum deviation (MD) a contour could deviate radially from the 'true' contour expressed as a fraction of each volunteer's mean myocardial width (MW). Significant MBF errors were avoided by setting tolerances of $MD \leq 0.4MW$, when considering the whole myocardium, $MD \leq 0.3MW$, when considering 6 radial segments, and $MD \leq 0.2MW$ for further subdivision into endo and epicardial regions, with the exception of the anteroseptal region, which required greater accuracy. None of the considered segmentation metrics correlated with MBF error, thus both segmentation metrics and MBF errors should be used to evaluate contouring algorithms.

Keywords Quantitative Myocardial Perfusion, Myocardial Localization, Segmentation, Evaluation

Submitted to: *Phys. Med. Biol.*

1. Introduction

Myocardial perfusion may be assessed by using dynamic contrast enhanced magnetic resonance imaging (DCE-MRI), in which injected gadolinium contrast agent is imaged passing through the heart in a dynamic image sequence. DCE-MRI is potentially superior to other currently available diagnostic imaging techniques for the assessment of myocardial perfusion due to its relative safety, compared to coronary artery disease assessment by X-ray angiography, and its high spatial resolution, compared to nuclear medicine, which enables sub-endocardial perfusion defects to be visualized (Wagner et al. 2003). However, DCE-MRI perfusion studies are currently only routinely assessed qualitatively, thus the benefits of the imaging modality are not fully utilized.

Myocardial blood flow (MBF) may be quantified by carrying out deconvolution analysis of signal intensity vs. time curves taken from regions of interest (ROI)s describing the left ventricular blood pool and myocardium (Jerosch-Herold et al. 2004). To obtain the curves these ROIs must be drawn for each frame in the DCE-MRI series. Manual contour drawing is time consuming and is a significant factor hindering the acceptance of quantitative perfusion into clinical practice (Jerosch-Herold et al. 2004). Understanding the level of accuracy required in the drawing of these contours is a key step in addressing this problem. How much a given error is likely to effect MBF measurements will dictate how carefully, and thus how quickly, a human contour drawer can perform their task. Such insights are also important in the evaluation of automated segmentation algorithms which tend to be evaluated with a wide range of segmentation error metrics making it difficult to cross compare algorithm performance (e.g. (Stegmann et al. 2005), (Adluru et al. 2006), (Santarelli et al. 2003)). The focus of such algorithms tends to be on accuracy of segmentation of the myocardial region of interest in high quality datasets, with poorer quality images being discarded as outliers (Santarelli et al. 2003). However poor quality images are a clinical reality and these algorithms may be aiming to achieve an unnecessary level of segmentation accuracy. Given the wide variety of sources of error in MBF estimates it may be the case that sacrificing segmentation accuracy in order to maintain robustness to poor image quality will have an insignificant effect on MBF estimates.

The purpose of this study was to investigate the relationship between geometrical error in myocardial segmentation and error in MBF. This relationship will provide a basis on which to decide acceptable error limits for myocardial contours, whether manually or automatically generated. It is also important for understanding whether automated segmentation algorithms evaluated in terms of segmentation metrics, e.g. (Stegmann et al. 2005), can be meaningfully compared with algorithms evaluated using MBF error, e.g. (Adluru et al. 2006), and may be instructive in deciding how best to evaluate such algorithms in the future.

2. Method

2.1. Datasets

Seventeen healthy volunteers (9 male, 8 females, mean age 34 years age range 24-48 years) with no history of heart disease, diabetes, hyperlipidaemia or chronic illness were recruited into the study. All volunteers had normal blood pressures and showed normal left ventricular mass as assessed by planimetry in short axis left ventricle stack images. Informed consent was taken from all volunteers in accordance with a study protocol approved by the regional ethics committee. All volunteers were instructed to refrain from caffeine for 24 hours prior to the examination.

2.2. Imaging

Myocardial perfusion DCE-MRI scans were performed on a 1.5T whole body MRI system (Philips Medical Intera systems, Best, the Netherlands). Volunteers were positioned supine with a flexible five element cardiac phased array receiver coil placed on their chest. Perfusion imaging was carried out with an intravenous injection of contrast (Magnevist, Schering, Berlin, Germany) at a dose of 0.05mmol/kg Gd-DTPA at a rate of 5ml/s followed by a 20ml saline flush via an automated infusion pump (Medrad Spectris Solaris, Medrad, Indianola, PA, USA). Volunteers were initially imaged under stress, which was pharmaceutically induced by an intravenous infusion of adenosine over 4 minutes at 140μg/min/kg. A rest scan was acquired approximately 15 minutes later. Mid-ventricular short axis DCE-MRI series were acquired at mid-systole and mid-diastole. The imaging sequence used has been previously described (Radjenovic et al. 2010). The images were acquired with a saturation recovery prepared single-shot gradient echo pulse sequence, TR/TE/α, 2.7ms/1.0ms/15°, partial Fourier = 0.67 (with missing data synthesized from acquired data using the conjugate symmetry of k-space), with two-fold SENSE (incorporating Constant LEvel AppeaRance (CLEAR: Philips Medical Systems, Best, The Netherlands) method for surface coil inhomogeneity correction) giving a total shot duration of 130ms, slice thickness 10mm, preparation pulse delay to the zeroth line of k-space of 150ms. The mean FOV was 332*mm* x 284*mm* (range: 290*mm* x 245*mm* to 410*mm* x 338*mm*). The mean scan pixel size was 2.27*mm* x 1.95*mm* (range: 2.08*mm* x 1.70*mm* to 2.64*mm* x 2.35*mm*). All images were reconstructed to a 256*x*256 matrix size giving a mean reconstructed pixel size in the image of 1.30*mm* x 1.11*mm* (range: 1.13*mm* x 0.96*mm* to 1.60*mm* x 1.32*mm*). Each dynamic image was obtained after an ECG triggering timed to image the heart in systole or diastole. Volunteers were instructed to hold their breath at end expiration, timed to coincide with the arrival of contrast into the heart, for as long as they were capable and then to resort to gentle breathing thereafter. This breath-holding strategy minimizes motion during the first pass of contrast agent through the myocardium, which is the data used for MBF estimation in this study.

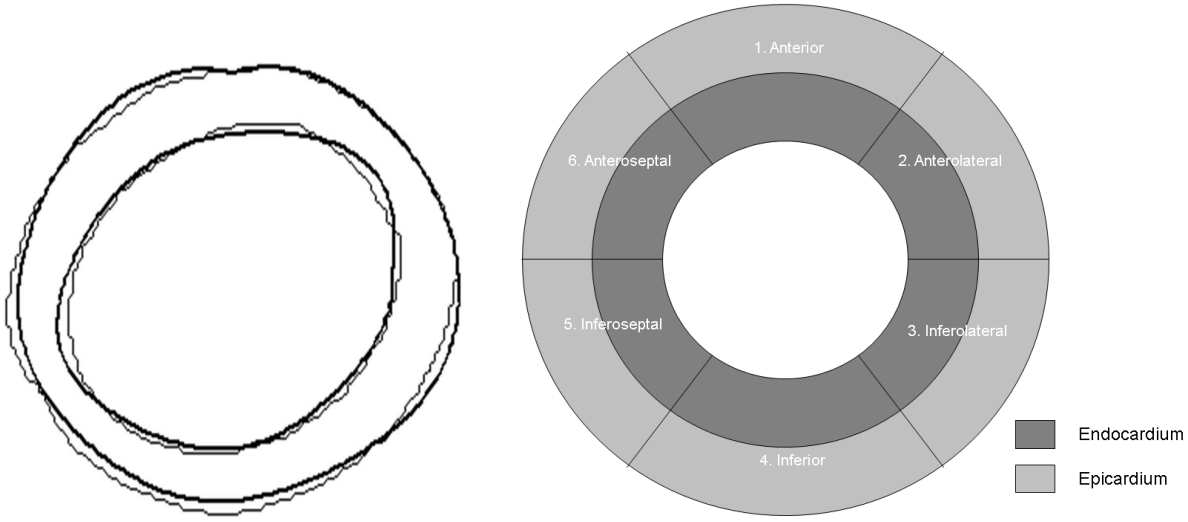
2.3. Contours

2.3.1. Manual Contouring Endocardial and epicardial contours were manually drawn by an expert user for every frame of each of the mid-systole cardiac DCE-MRI dynamic series using dedicated cardiac image analysis software (Mass 5.0, Medis, Leiden University, Leiden, The Netherlands). A further contour was drawn within the left ventricular blood pool, avoiding the papillary muscles (see figure 3). In a given dynamic series the image exhibiting maximum contrast between the myocardium and surrounding tissues was used to draw endocardial and epicardial contours and a region in the left ventricular blood pool, avoiding the papillary muscles. These contours were then copied to the full time series and manually translated to compensate for breathing motion.

To ascertain whether our simulated contour errors were representative of human contour errors a subset of 11 patients were contoured a second time by the same manual contourer to assess intra-observer variability. A second contourer also contoured the dataset to assess inter-observer variability. To measure the error between each contour and the reference contour the geometric centre of gravity of the two circular contours were aligned and the Euclidean distance between each point on the reference contour and the nearest point on the contour being investigated was calculated. The distribution of these errors was then compared to the difference between the reference contour and the random contour errors simulated in this paper using the Kolmogorov-Smirnoff test (Massey 1951), which tests whether two samples have been drawn from the same underlying continuous population. The inter/intra observer variability data was analyzed considering stress and rest and endo and epi contours separately and then finally analyzed considering all cases together.

The myocardium was divided into 6 circumferentially equidistant partitions according to the model proposed by the American Heart Association (AHA) (Cerqueira et al. 2002) for the mid ventricular slice. The regions were then further divided circumferentially into endocardial and epicardial compartments resulting in a total of 12 partitions as illustrated in figure 1b.

2.3.2. Random Contour Errors Contour errors were generated by introducing random radial deviations into the manual contour. The manual contour was represented as a circular spline by automatically placing equally spaced knot points along the defined contour. The knot points of the spline were offset by a random displacement allowed to range between $\pm MD$, where MD is the maximum deviation , being the maximum distance the contour may deviate from the 'true' contour. MD was expressed as a fraction of the mean myocardial width (MW) so that the degree of contour error was normalized to the size of the heart. A new contour was then generated from these offset knot values to represent the erroneous curve. Figure 1a shows an example of a generated erroneous contour with $MD = 0.1MW$. We chose to use ten knot points to represent the contour to maintain a realistically smooth, approximately circular contour. Each randomly generated set of offset values was applied to contours at all time points in all patients. This process was repeated over 30 iterations of the random offset value generation. The process was repeated using $MD = 0.1MW, 0.2MW, 0.3MW, 0.4MW$ and $0.5MW$. T-tests and F-tests between the 'true' and erroneous contour datasets were



(a) (b)

Figure 1: (Left) Example of generated contour errors. The thin line shows the manual contour and the bold line shows the generated erroneous contour generated with a maximum deviation (MD) of $0.1MW$. (Right) Schematic showing the partitioning of the myocardial ROI.

performed to test for significant differences in the means and variances of the distributions respectively.

2.3.3. Systematic Contour Errors Systematic under- and overestimates of the myocardial region of interest were generated by setting the MD to a constant offset value over all the knot points. For each frame of the cardiac DCE-MRI perfusion dataset the myocardial contours were modified by systematic MD values ranging from $-0.5MW$ to $+0.5MW$. Differences were calculated, at rest and stress, between the MBFs estimated from modified and unmodified contours and the difference in the means of the resulting distributions was assessed using a statistical t-test.

2.4. Segmentation metrics

We evaluated our segmentation errors with two geometric segmentation metrics, Hausdorff distance (HD) (Beauchemin 1998), based on the distance between the two contours and Dice's similarity coefficient (DSC) (Dice 1945), (Yasnoff et al. 1977), based on the overlapping areas of the two contours and in terms of error in myocardial blood flow (MBF).

2.5. Quantitation of MBF

The MBF was quantified from the left ventricular blood pool and myocardial tissue signal intensity vs. time curves using a Fermi constrained deconvolution method (Jerosch-Herold et al. 1998), summarized as follows. The amount of tracer in a stationary, linear system,

represented by the tissue curve $q_{\mathrm{myo}}(t)$, can be related to the tracer concentration at the inlet, the arterial input function (AIF), represented by the blood pool curve $c_{\mathrm{b}}(t)$, convolved with the impulse response function $F \cdot R(t)$:

$$q_{\mathrm{myo}}(t) = F \cdot R(t) \otimes c_{\mathrm{b}}(t) \tag{1}$$

$R(t)$ is the residue function, which represents the fraction of tracer that remains in the myocardium at time t and F is the rate of flow. At time zero $R(t)$ will be one therefore $F \cdot R(t) = F$, thus by establishing $F \cdot R(t)$ we can obtain an estimate for the flow. Deconvolution of equation 1 is an ill-posed problem but an estimate for $F \cdot R(t)$ can be obtained by constraining the deconvolution operation by modeling the impulse response function with a Fermi function:

$$R(t) = \frac{\left(1 + exp(-\frac{\omega}{\tau})\right)}{\left(1 + exp(-\frac{t-\omega}{\tau})\right)} \tag{2}$$

ω and τ are the parameters of the model and do not have a direct physiological interpretation. Using a least squares fitting approach the parameters of equation 2 can be optimized to fit the observed data and establish a best estimate for the impulse response function and thus the flow.

2.6. Saturation Correction

For contrast agent (CA) doses of 0.05mmol/kg MRI signal intensity does not vary linearly with concentration. Signal intensity begins to saturate at higher CA concentrations causing a blunting of the AIF peak and a subsequent over estimate of MBF. We used the method described by (Larsson et al. 1996) and validated by (Fritz-Hansen et al. 1996), (Fritz-Hansen et al. 2008) to convert signal intensity to concentration, thereby correcting for this signal saturation. The fundamental assumption (Rohrer et al. 2005) is that change in longitudinal relaxation rate T_1 due to a given concentration $c(t)$ of contrast agent at time t can be related as follows:

$$\left(\frac{1}{T_1(t)} - \frac{1}{T_0(0)}\right) = r_1 c(t) \tag{3}$$

where $T_1(0)$ is the relaxation time without the CA, $T_1(t)$ is the relaxation time with the CA and r_1 is the CA relaxivity. Thus the concentration $c(t)$ of CA at time t can be expressed as:

$$c(t) = \frac{\Delta R_1(t)}{r_1} \tag{4}$$

where $\Delta R_1(t) = \left(\frac{1}{T_1(t)} - \frac{1}{T_1(0)}\right)$. Signal intensity and T_1 are related by the MR signal equation for the saturation recovery prepared single-shot gradient echo pulse sequence as follows :

$$S = \Psi \cdot \left[(1 - e^{-PD \cdot R_1})a^{n-1} + b\frac{1 - a^{n-1}}{1 - a}\right] \tag{5}$$

(Equation 5 is adapted from (Larsson et al. 1996) where it is derived for the inversion recovery turbo-FLASH sequence. The modification for saturation recovery is trivial.) PD is the pre-pulse delay between the saturation pulse and the central line of k-space, n is the number of applied pulses of flip angle α to the central line of k-space, $a = cos(\alpha)e^{-TR \cdot R_1}$, $b = 1 - e^{-TR \cdot R_1}$. Ψ is a calibration constant dependent on receiver gain, proton density and α. (Larsson et al. 1996) used a pre-contrast T_1, measurement to calculate Ψ from equation 5. Ψ is then assumed to be constant throughout out the dynamic acquisition thus T_1 values for images post-contrast agent arrival can be calculated from equation 5 which can then be used to calculated CA concentration from equation 4. An analytical solution to equation 5 for R_1 is not possible so it is solved using numerical methods. In our implementation a single-variable nonlinear zero finding algorithm was used, (fzero.m Maltlab7 R2009b) (Brent 1973). We used an assumed T_1 from a weighted average of published blood T_1 measurements as opposed to a bespoke, patient specific T_1 measurement to calculate Ψ. By assuming that the calculated calibration constant Ψ was the same in the blood and the myocardium the signal to concentration conversion process described was also applied to the myocardial signal intensity vs. time curve as well. Figure 2 illustrates the effects of this correction method on a typical AIF at stress.

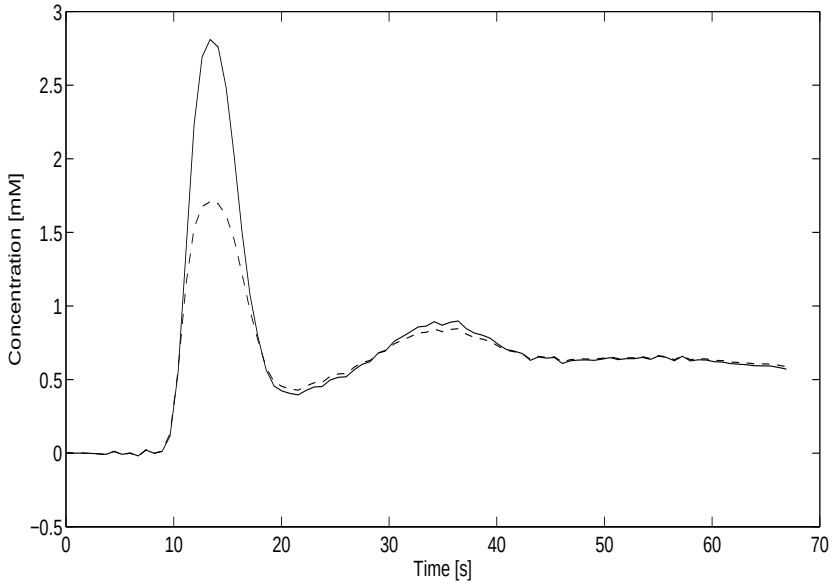


Figure 2: Example arterial input function (AIF) concentration vs. time curves showing original (dashed line) and saturation corrected (solid line) at stress. NOTE: As signal intensity and concentration are not linearly related it is not strictly possible to show both curves on the same axis. Figure 2 was generated by fitting the tail of the SI curve to the concentration curve.

2.7. Data preparation

Both myocardial and AIF curves were interpolated using piecewise Hermitian interpolation (Fritsch & Carlson 1980) to increase the apparent temporal resolution and to ensure equal

temporal spacing between all points. To correct for the time delay between the onset of contrast in the AIF and the myocardium the Fermi constrained deconvolution operation was performed iteratively over a range of time increment steps and the time step yielding the best χ^2 fit to the data was used. Baseline correction by subtraction of the mean pre-contrast signal intensity was necessary for the resting curves due to remnant Gadolinium in the blood from the stress study. The conversion to concentration step automatically generates zero mean baselines in the stress dataset. The bolus arrival time was obtained using the piecewise continuous regression model of (Cheong & et al. 2003), which fits a combination of two straight lines to the curve. All data points prior to the resulting bolus arrival time were then classed as pre-contrast signal. Data was limited to the first-pass of the CA through the heart, which was identified as the first valley in the AIF after the CA entered the LV. This was automatically detected by identifying the downslope of the first peak in the blood curve by finding the minimum of the first differential of the smoothed blood curve. The end of the first-pass was then taken as the next point where the differentiated curve crossed the the line $x = 0$.

2.8. Implementation

The above method was implemented in MATLAB (The Mathworks, Natick, MA), using an assumed blood T_1 value of $1435ms$, which was derived from an average (weighted for study population) of the measurements published in (Flacke et al. 2001)(Klein et al. 2004)(Messroghli et al. 2004)(Sharma et al. 2006). The conversion from signal intensity to concentration was successful in 16/17 volunteers with one case failing because no T_1 value could be found to solve the signal equation. This volunteer was excluded from the study.

3. Results

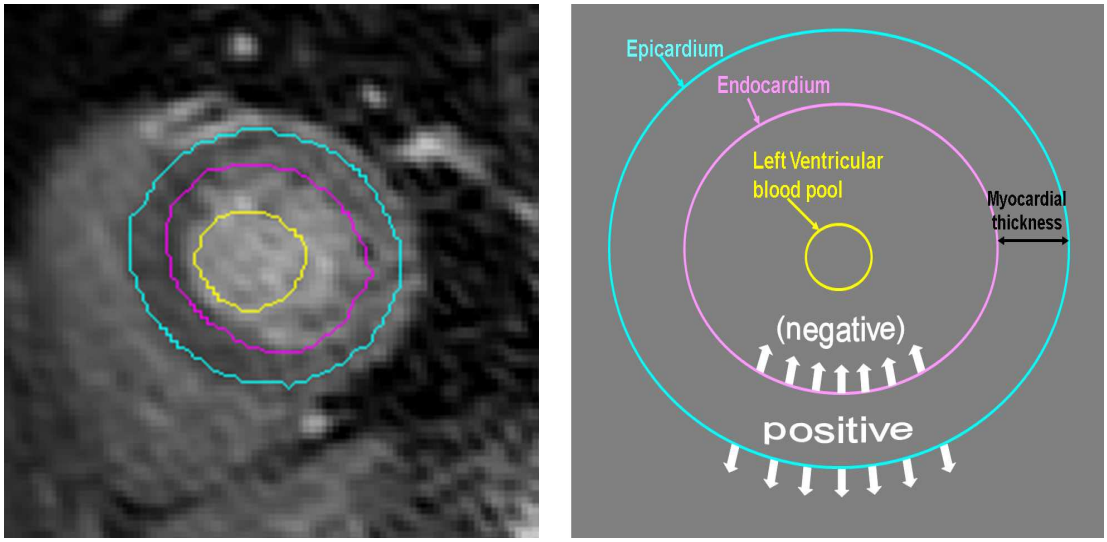
Using the manual contours the mean (± standard deviation) MBF at rest was $1.24 \pm 0.35ml/g/min$ and at stress was $3.48 \pm 0.67ml/g/min$. The mean myocardial width (MW) was 5.8 voxels (range: 3.3 voxels to 8.6 voxels) and $6.9mm$ (range: $4.2mm$ to $10.3mm$).

3.1. Segmentation Metrics

Table 1 shows Pearson's correlation scores between MBF error and maximum deviation (MD), Dice's similarity coefficient (DSC) and Hausdorff distance (HD) for endo and epicardial modification at rest and stress considered separately and as one data set. None of the segmentation metrics correlated with MBF error with the most significant correlation at $r = -0.32$.

3.2. Systematic Contour Errors

Figure 4 shows the spread of MBF errors for each MD for the endo and epicardium at rest and stress over the entire myocardium. Positive MD values correspond to contours



(a) (b)

Figure 3: (Left) A single frame of a cardiac perfusion sequence showing manual contouring. The left ventricular cavity is filled with Gd-DTPA and appears bright against the surrounding myocardium. (Right) A schematic representation of the systematic contour errors (right) illustrating the directions of the positive and negative errors.

Table 1: Table of the Pearson's correlation r-value between MBF error [ml/g/min] and each of the three segmentation error metrics: Maximum Deviation (MD) expressed as a fraction of the mean myocardial width (MW), Dice's similarity coefficient (DSC) [no units] and Hausdorff distance (HD) [mm]. Results are shown considering errors in rest/stress and endocontour/epicontour separately and finally over all data.

Contour Error	**MD**	**DSC**	**HD [mm]**
Rest Endocardium	0.03	0.03	0.02
Rest Epicardium	-0.14	0.15	-0.17
Stress Endocardium	0.07	-0.10	0.04
Stress Epicardium	-0.22	0.26	-0.32
Rest and Stress, Endo- and Epi-contour	-0.08	0.17	-0.13

modified circumferentially outwards (moving away from the centre of the myocardial circle) and negative contours correspond to contours modified circumferentially inwards, (see figure 3). Errors in MBF were calculated as the difference between the MBF estimated with the modified contours and the MBF estimated with the manual contours. Student's t-test between the modified and unmodified MBF error populations yielded non significant p-values for all MD values.

Figure 5 shows the mean MBF errors for the 6 separate myocardial segments. Individual t-tests for each segment showed no significant difference in mean MBF except for the inferoseptal segment, where a resting epicardial MD of $0.5MW$ gave ($p = 0.05$) and a stress endocardial MD of $-0.5MW$ gave ($p = 0.02$).

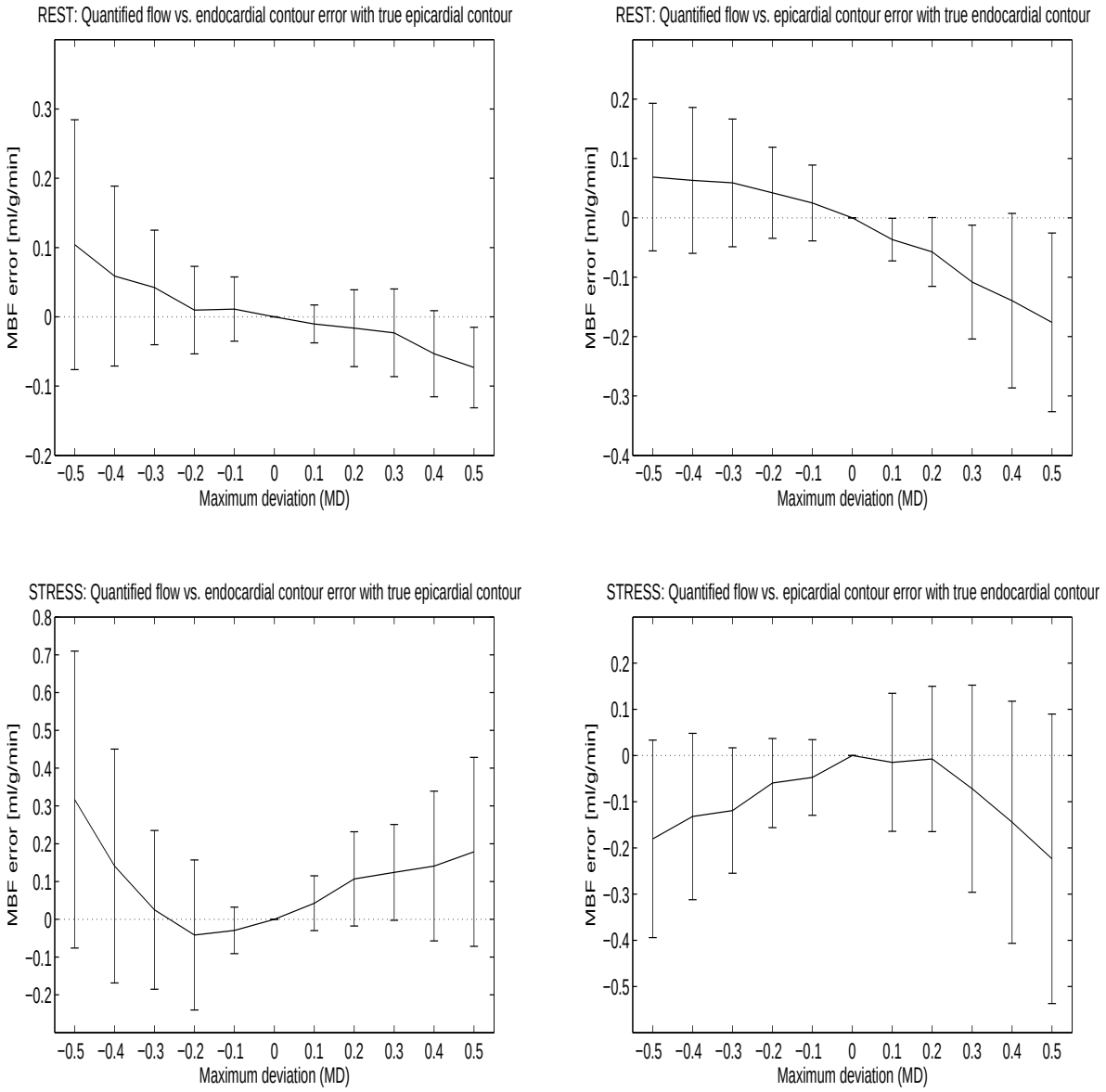


Figure 4: Mean MBF errors vs. MD (expressed as a fraction of the mean myocardial width) for systematic contour errors in the rest endocardial contour (top left) rest epicardial contour (top right), stress endocardial contour (bottom left) and stress epicardial contour (bottom right). Error bars show the standard deviations.

Figure 6 shows the corresponding analysis when the endocardium and epicardium were considered separately. When the epicardial contour is modified only signal from the epicardium is incorporated in the analysis. When the endocardial contour is modified only the endocardial tissue is considered. The t-test for sub-myocardial segments showed generally more statistically significant results than for transmural segments. Significant differences were seen in the endocardial inferospetal segment for $MD = -0.5MW$ at rest ($p < 0.05$) and stress ($p < 0.03$), the epicardial inferoseptal segment at stress for $MD = 0.5MW$ ($p = 0.01$) and $MD = 0.4MW$ ($p = 0.04$), the epicardial inferior segment at rest for $MD = 0.5MW$

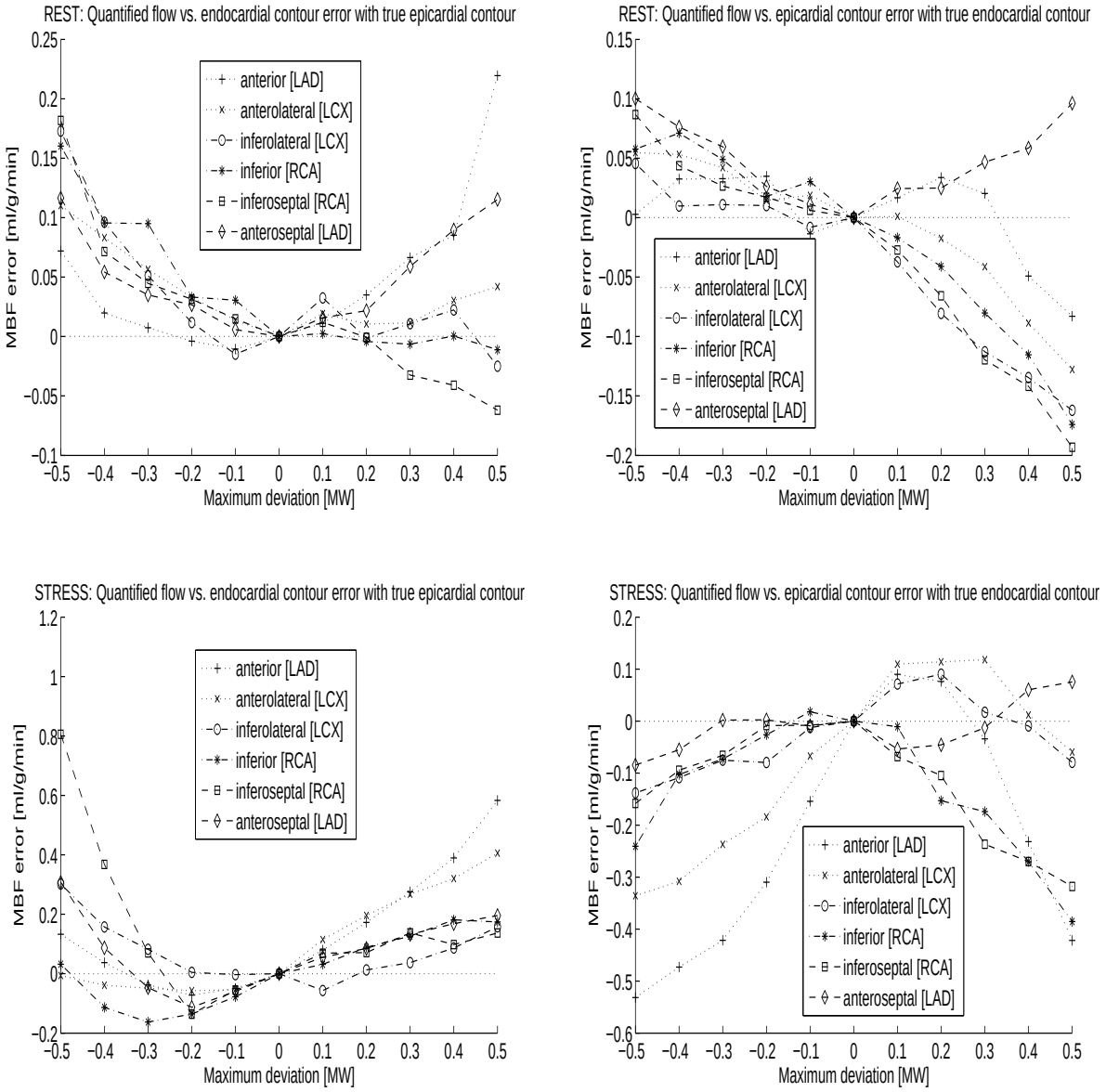


Figure 5: Mean MBF error vs. MD (expressed as a fraction of the mean myocardial width) for systematic contour errors in the rest endocardial contour (top left) rest epicardial contour (top right), stress endocardial contour (bottom left) and stress epicardial contour (bottom right).

($p = 0.03$), the epicardial anterior segment for $MD = 0.5MW$ at rest ($p = 0.04$) and stress ($p = 0.03$) and in the epicardial inferior segment for $MD = 0.5MW$ at rest ($p = 0.02$) and stress ($p = 0.02$).

3.3. Random Contour Errors

3.3.1. Mean Myocardium Figure 7 shows the effect of the random contour errors on MBFs estimated over the entire myocardium. Each box-plot represents MBF errors incurred using contours whose random deviations were limited to the given MD on the x-axis. The central

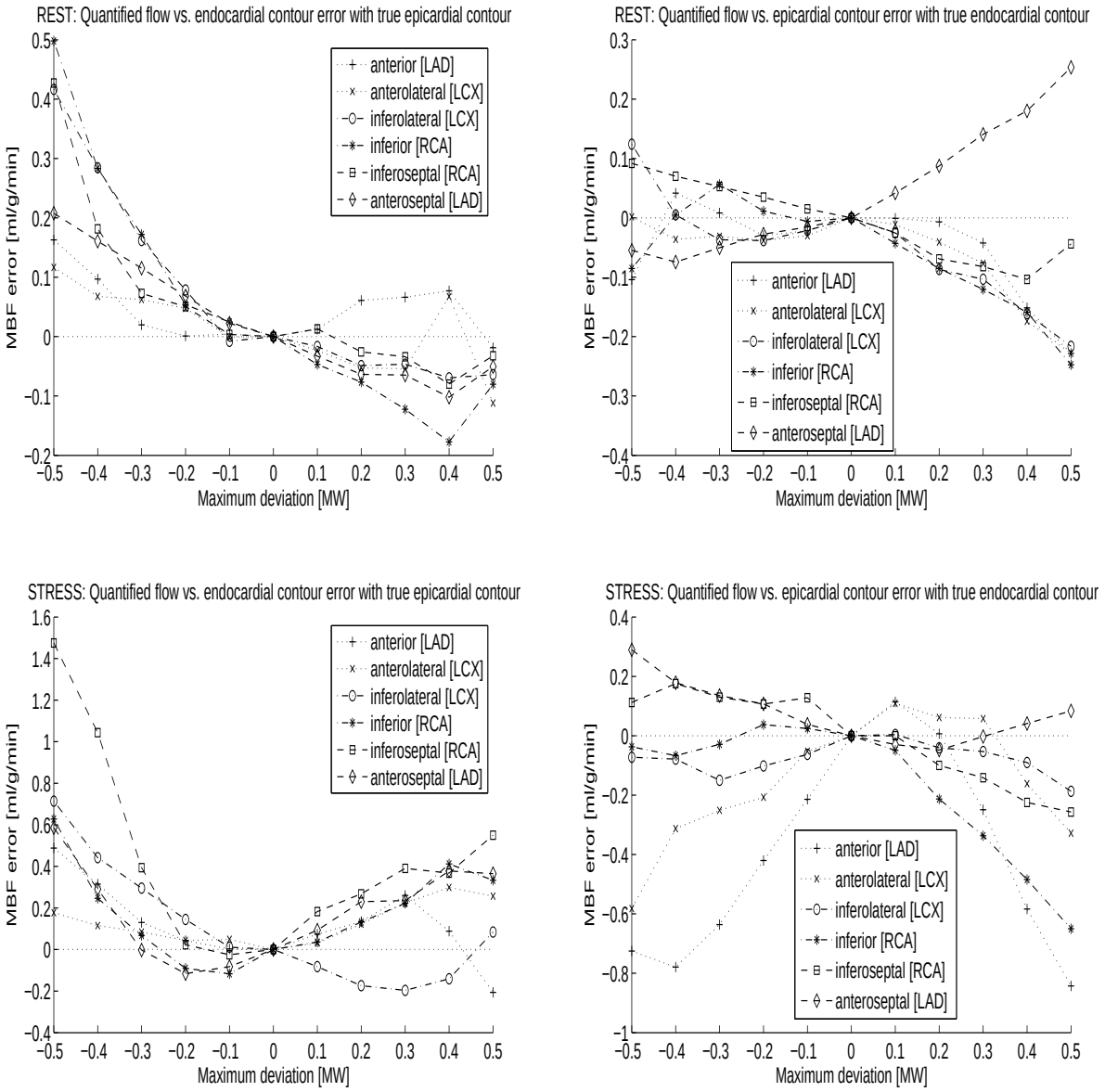


Figure 6: Segmental mean MBF errors divided into endocardial and epicardial segments vs. MD (expressed as a fraction of the mean myocardial width) for systematic contour errors in the rest endocardial contour (top left) rest epicardial contour (top right), stress endocardial contour (bottom left) and stress epicardial contour (bottom right).

line, box and whiskers correspond to the median, interquartile range and 95% percentile range respectively. F-tests and t-test for differences in variance and mean MBF error between manual and modified contours were non-significant ($p > 0.05$) in all cases except a MD of $0.5MW$ in the resting epicardium (t-test: $p = 0.03$). At stress an MD of $0.5MW$ in the myocardium approached statistical significance (t-test: $p = 0.07$).

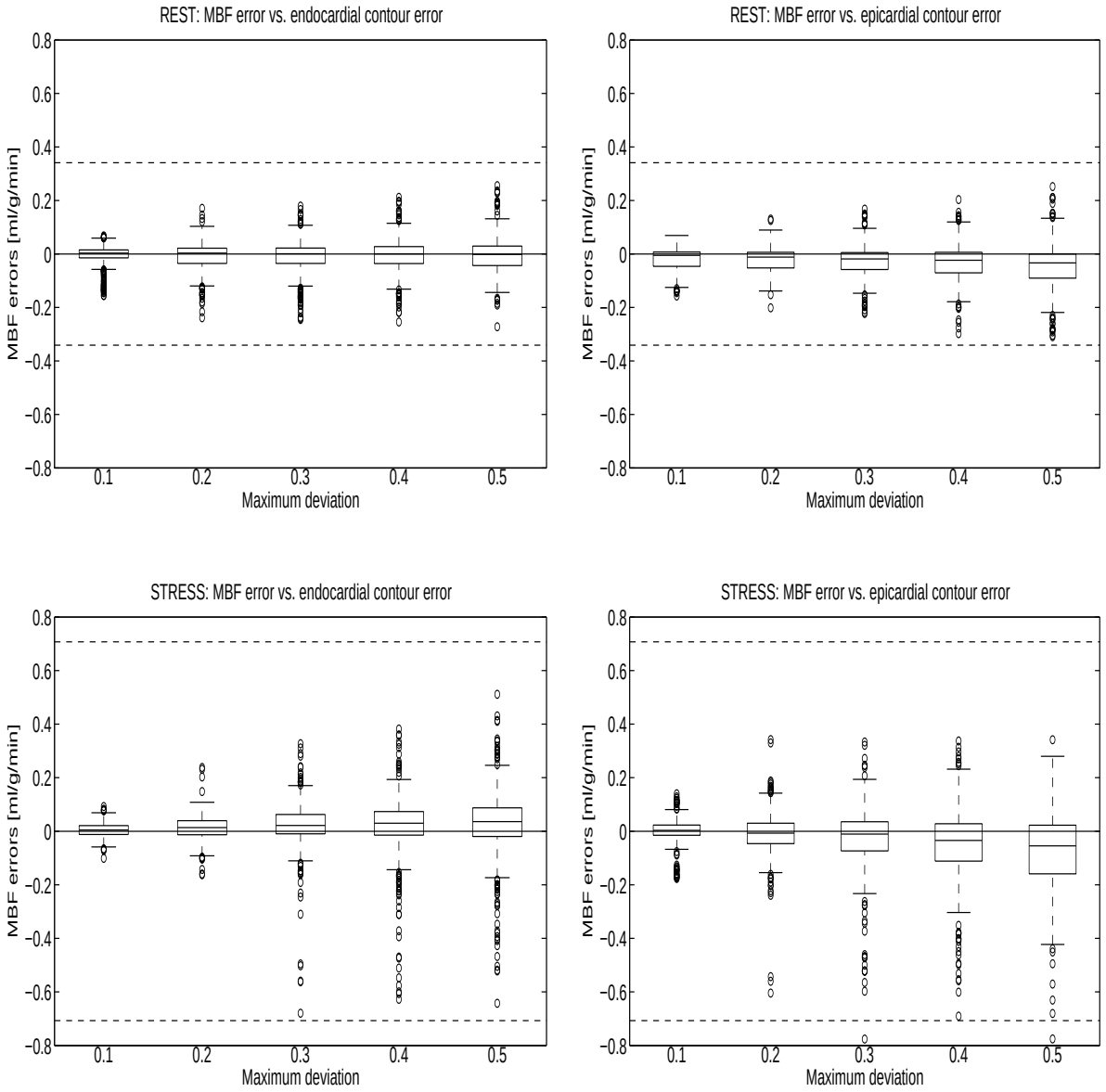


Figure 7: Box plots of MBF errors vs. MD (expressed as a fraction of the mean myocardial width) in the rest endocardial contour (top left) rest epicardial contour (top right), stress endocardial contour (bottom left) and stress epicardial contour (bottom right). Lines, box edges and whiskers of each box plot correspond to the median, inter-quartile range and 95% percentile range of MBF errors respectively. The dashed black lines depict ± one standard deviation of the MBF values obtained with the manual contours.

3.3.2. Segmental Myocardium The effects of random contour errors on MBF in the 6 separate myocardial segments showed similar trends to figure 7, with the spread of MBF errors increasing with MD. Statistically significant results are shown in table 2.

3.3.3. Endo and Epicardium Considering the endocardium and epicardium as two separate regions showed similar trends with a more dramatic increase in the spread of MBF values with

Table 2: Table of statistically significant MBF errors generated by random contour errors considering the six myocardial segments. MD corresponds to the maximum deviation at which MBF errors became significant. Cases where significant ($p \leq 0.05$) differences were not observed are not shown.

Segment	Rest/Stress	Endo/Epi	MD[MW]	Test	p-value
Anterior (LAD)	Stress	Endo	0.5	F-test	0.04
Anterior (LAD)	Stress	Epi	0.5	t-test	0.04
Anterolateral [LCX]					N.S.
Inferolateral [LCX]	Stress	Endo	0.5	F-test	< 0.01
Inferior [RCA]	Rest	Endo	0.5	F-test	0.05
Inferior [RCA]	Rest	Epi	0.4	t-test	0.04
Inferior [RCA]	Stress	Epi	0.5	t-test	0.02
Inferoseptal [RCA]	Stress	Epi	0.5	t-test	0.02
Anteroseptal [LAD]	Rest	Epi	0.5	t-test	0.02
Anteroseptal [LAD]	Rest	Epi	0.2	F-test	< 0.01

increasing MD. All statistically significant differences in mean (t-test) and variance (F-test) of MBF errors are reported in table 3.

3.4. Inter and intra observer variability

Figure 8a shows the distribution of contour errors between the manually drawn contours of the first and second observers (inter-observer) and the distribution of random contour errors between the simulated and manual contours for each MD. Kolmogorov-Smirnoff tests between each simulated distribution and the inter-observer distribution yielded the following p-values: $0.1MW(p = 0.01)$, $0.2MW(p = 0.08)$, $0.3MW(p = 0.08)$, $0.4MW(p = 0.38)$ and $0.5MW(p = 0.93)$. Figure 8b shows the corresponding distribution for the repeated manually drawn contours (intra-observer). Kolmogorov-Smirnoff tests between the simulated distributions and the intra-observer distribution yielded the following p-values: $0.1MW(p = 0.03)$, $0.2MW(p = 0.19)$, $0.3MW(p = 0.19)$, $0.4MW(p = 0.67)$ and $0.5MW(p = 0.93)$. Figure 8 includes contour errors from stress and rest and endo and epi contours all together. Separate analyses of each of these four cases was carried out yielding non significant ($p < 0.05$) Kolmogorov-Smirnoff tests in all cases.

4. Discussion

Using the manual contours the mean (± standard deviation) MBF at rest was $1.24 \pm 0.35 ml/g/min$ and at stress was $3.48 \pm 0.67 ml/g/min$, which is consistent with other studies measuring MBF in healthy volunteers ((Pack et al. 2008), (Jerosch-Herold et al. 1999)).

4.1. Segmentation Metrics

None of the segmentation metrics considered correlated with MBF error for the random contour error simulation. In the light of the results from the systematic contour error

Table 3: Table of statistically significant MBF errors generated by random contour errors considering the endocardium and epicardium separately. MD corresponds to the maximum deviation at which MBF errors became significant. Cases where significant ($p \leq 0.05$) differences were not observed are not shown.

Segment	Rest/Stress	Endo/Epi	MD[MW]	Test	p-value
Anterior (LAD)	Stress	Epi	0.5	t-test	< 0.01
Anterior (LAD)	Rest	Epi	0.4	F-test	0.02
Anterior (LAD)	Stress	Endo	0.4	F-test	0.02
Anterior (LAD)	Rest	Endo	0.5	F-test	0.03
Anterolateral [LCX]	Rest	Endo	0.4	F-test	0.02
Anterolateral [LCX]	Stress	Epi	0.4	F-test	0.04
Inferolateral [LCX]	Rest	Endo	0.3	t-test	0.04
Inferolateral [LCX]	Rest	Endo	0.3	F-test	0.05
Inferolateral [LCX]	Stress	Endo	0.4	F-test	< 0.01
Inferior [RCA]	Stress	Epi	0.4	t-test	0.02
Inferior [RCA]	Rest	Endo	0.5	t-test	0.01
Inferior [RCA]	Rest	Epi	0.5	t-test	0.05
Inferior [RCA]	Stress	Endo	0.5	t-test	0.01
Inferior [RCA]	Stress	Endo	0.3	F-test	0.04
Inferior [RCA]	Rest	Endo	0.5	F-test	< 0.01
Inferoseptal [RCA]	Stress	Endo	0.3	t-test	0.05
Inferoseptal [RCA]	Stress	Endo	0.4	F-test	0.04
Anteroseptal [LAD]	Rest	Epi	0.2	t-test	0.02
Anteroseptal [LAD]	Stress	Endo	0.5	t-test	0.01
Anteroseptal [LAD]	Rest	Epi	0.1	F-test	0.01
Anteroseptal [LAD]	Stress	Endo	0.3	F-test	0.01
Mean	Rest	Epi	0.5	t-test	0.03
Mean	Stress	Endo	0.5	t-test	0.05
Mean	Stress	Endo	0.4	F-test	0.04
Mean	Rest	Endo	0.5	F-test	0.01

simulations this result is not surprising. A given contour error measured by DSC or HD may correspond to a movement of either the endocardial or epicardial contour into any of a variety of surrounding tissues with conflicting effects on MBF. The conclusion is that neither MBF error or segmentation alone is an adequate measure of contour error as there are too many conflicting factors affecting the relationship between these two measures. Thus if contour errors are random in nature then analysis in terms of MBF error may only show an increase in the variance of the MBF errors, with insignificant changes in the mean MBF. This could lead to misleading claims about the accuracy of an automated algorithm. If the algorithm produces contour errors of a systematic nature then measures of MBF error may be correlated with segmentation error, as described in figure 4. However these relationships are not linear and it is unlikely that an algorithm would induce systematic errors as uniformly as those simulated here. Therefore contouring algorithms for DCE-MRI myocardial perfusion should ideally be evaluated by both geometric segmentation metrics and in terms of MBF.

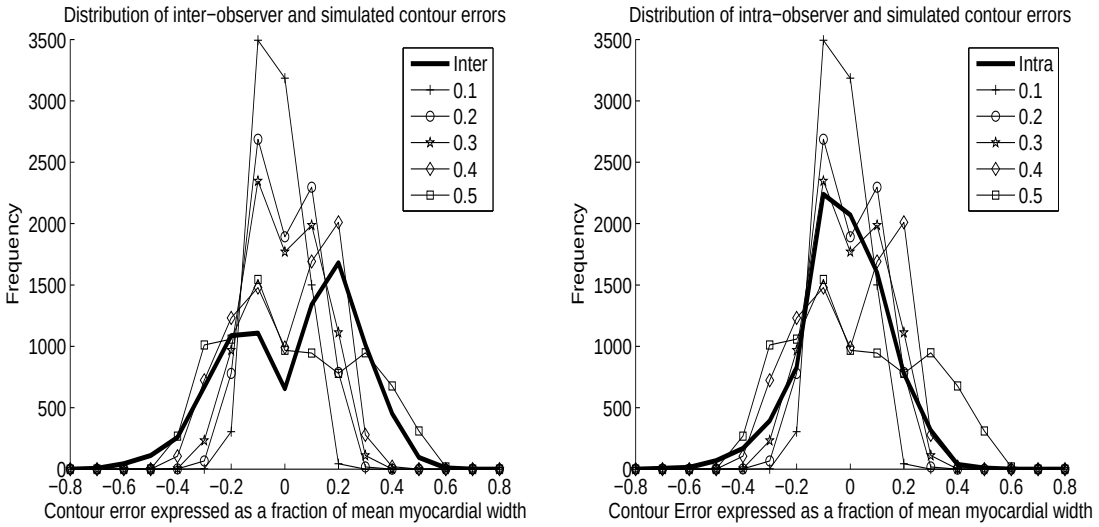


Figure 8: Distribution of inter (a) and intra (b) observer contour errors (thick lines). The corresponding distributions of contour errors between manual and simulated contours at each MD setting are also shown (thin lines)

4.2. Systematic Contour Errors

The application of systematic contour errors to the dataset is useful for understanding how MBF errors vary depending on which contour (endocontour/epicontour) has moved and which direction it has moved in. They also serve to simulate systematic conservative or generous contouring. A given contourer may be overly concerned with avoiding non-myocardial voxels or conversely including all myocardial voxels thereby making this type of systematic error. It is conceivable that such global contours could occur with automated contouring algorithms as well. For example active contour based methods with non-ideal stopping functions may generate consistent global over or underestimates in the contour, and an active appearance model driven method such as (Stegmann et al. 2005) will be as conservative as the manual data set on which it is trained.

4.2.1. Mean Myocardium Systematic trends in MBF error were seen as the contours were modified. These trends can be explained in the light of previously observed physiological flow properties of the myocardium. Animal studies have measured the presence of a transmural MBF gradient across the resting myocardium favoring the endocardium, which was no longer observed under stress conditions (Ball & Bache 1975), (Parks & Manohar 1983). If these properties remain true in the human myocardium then the trends in figure 4 can be explained as follows:

Variation of endocardial contour at rest. The MBF error becomes more positive for negative contour errors as the endocardial contour encroaches on the voxels within the left ventricular blood pool. There is rapid signal enhancement in the left ventricular cavity

thus these voxels, erroneously incorporated into the myocardial region, will lead to MBF overestimation (i.e. positive MBF errors). For positive contour errors the higher MBF endocardial voxels become excluded thus the relative flow goes down and the MBF error becomes negative.

Variation of epicardial contour at rest. For negative contour errors the MBF error becomes more positive as the relatively low MBF epicardial voxels are excluded from the myocardial region. Positive contour errors will incorporate non-myocardial voxels of zero signal enhancement into the region thus reducing MBF and causing negative MBF errors.

Variation of endocardial contour at stress. As at rest the negative contour errors increase the MBF error. At MD values above $-0.2MW$ this effect is not apparent, which may be due to conservative contouring by the manual contourer. There is a clear increase in MBF error with positive endocontour errors which implies a transmural myocardial gradient at stress with the epicardium more highly perfused than the endocardium. As the endocardial contour encroaches on the myocardium the low MBF voxels in the endocardium are excluded, thus increasing the MBF error.

Variation of the epicardial contour at rest. As at rest positive contour errors reduce MBF error. There is also a reduction in MBF error with negative contour error which is further evidence for a transmural flow gradient at stress causing negative MBF errors as the higher flow voxels in the epicardium are excluded.

Whereas the presence of a resting transmural gradient is accepted there is conflicting evidence for the presence of a transmural flow gradient at stress. (Lee et al. 2004) observed no such gradient in healthy myocardial tissue at stress, but (Christian et al. 2004) observed a transmural (epi>endo) gradient at stress that was statistically significant as did (Radjenovic et al. 2010) in systole. However, the effect may in part be due to measurement errors inherent in the acquisition or analysis. A possible explanation might be the inclusion of endocardial dark rim artefacts in the myocardial region of interest, which could null the endocardial MBF values thereby generating the observed gradient.

Considering the whole myocardial region systematic contour errors of up to half the mean myocardial width did not yield statistically significant errors in MBF, figure 4.

4.2.2. Segmental Myocardium A measurement of the global MBF is of limited use in investigating coronary artery disease, which induces localized flow defects. The American Heart Association (AHA) model (Cerqueira et al. 2002) partitions the mid-myocardial slice into 6 circumferentially equidistant partitions that are associated with specific coronary arteries enabling the link between the perfusion imaging result and the required intervention. The transmural variation in MBF between endocardial and epicardial tissue has also been shown to be related to arterial stenosis, (Lee et al. 2004), therefore it is important to consider the effect of contour errors on these partitions. For these reasons our myocardial regions were also divided into the 12 partitions illustrated in figure 1b.

In general the 6 segment curves followed the same general trends as those for the whole myocardium with some notable exceptions. At rest positive epicardial contour errors for the anteroseptal segment yielded positive MBF errors, opposing the general trend of negative

errors. This is due to inclusion of blood voxels in the right ventricle directly adjacent to this myocardial segment, which exhibit rapid signal enhancement.. The effect is much less apparent on the corresponding stress plot because the relative effect of the right ventricular voxels is reduced with respect to the higher genuine myocardial MBF at stress. At stress the contour errors pushing the contours inside the myocardium (i.e. negative epicontour errors and positive endocontour errors) appear to effect the anterior segments more profoundly than the remaining segments. There is no reason to expect a stronger transmural gradient in the anterior myocardium so this observation is unexplained. In general the effects explained above in terms of transmural gradients for the mean myocardial data are obscured, either by errors in the measurement process, which is to be expected due to poorer SNR in the smaller segments, or by genuine heterogeneity of MBF gradients around the myocardium. Segmental analysis of systematic errors, figure 5, showed non-significant MBF errors except for the inferoseptal segment where a resting epicardial MD of $0.5MW$ gave ($p = 0.05$) and a stress endocardial MD of $-0.5MW$ gave ($p = 0.02$), thus a tolerance level of $MD = 0.4MW$ would avoid significant errors in MBF.

4.2.3. Endo and epicardium When the endocardial and epicardial layers were considered separately significant errors in more segments at $MD = \pm 0.5MW$ were seen. This is due to the greater percentage effect of a given voxel offset on the ROI. A tolerance of $MD = 0.3MW$ is required to avoid all significant MBF errors with the most susceptible region being the epicardial inferior segment at rest which exhibits significant MBF errors at $MD = 0.4MW$ (p=0.04).

These results suggest that, even in the unlikely event of a systematic error effecting the entire contour, errors limited to an MD of 0.3 times the mean myocardial width will not incur a statistically significant MBF error. Generally the largest MBF errors were seen when the contours passed outside of the myocardium, either epicardially or endocardially, thus conservative contouring is preferable to generous contouring for healthy volunteers. Where possible, segmentation algorithms should err on the side of placing the endocardial contour within the myocardium, however, in ischaemic patients with subendocardial abnormalities the placement of the subendocardial contour may be more critical than these results suggest for diagnosing ischaemia.

4.3. Random Contour Errors

4.3.1. Mean Myocardium The box-plots in figure 7 illustrate how the spread of MBF errors increases as the maximum allowed voxel offset increases. For a random contour error applied to a linear MBF error space one would expect the median MBF error to remain at zero independent of the size of the contour error. However figure 4 illustrates the fact that the MBF error space is not linear. For this reason as the maximum random error increases the median MBF error deviates from zero in figure 7. However contour errors up to a MD of $0.4MW$ did not cause a statistical shift in mean MBF. The increase in spread of MBF values with contour error is expected but F-tests did not show this to be significant even at 0.5 of

the mean myocardial width. The 95% confidence interval for MBF errors did not exceed the one standard deviation line of 'true' MBFs up to $MD = 0.5MW$. This shows that the contour errors simulated did not induce a statistically significant change in the distribution of MBFs implying that the variance induced in MBF estimates from our simulated contour errors is not significant compared to the natural variance of MBF within the healthy population. A statistically significant effect was seen for $MD = 0.5MW$ in the resting epicardium (t-test: $p = 0.03$). This is predominantly caused by the effect of the epicardial contour bleeding into the right ventricular blood pool as is clarified by the segmental analysis. These results suggest that a safety tolerance of $MD = 0.4MW$ would be acceptable for the anlaysis of MBF in the whole myocardium.

4.3.2. Segmental Myocardium Segmental analysis of the myocardium yielded similar trends in the spread of MBF errors to figure 7 with the spread in MBF error increasing more severely with MD than in the mean myocardium due to the more profound effect a given contour error has on smaller ROIs. In general statistically significant MBF errors can be avoided by setting a tolerance of $MD \leq 0.3MW$, (table 2). The exception is the anteroseptal segment for which statistically significant changes in the variance of the MBF error distributions were seen for the resting epicardium for $MD = 0.2MW$ (F-test: $p < 0.01$). Bleeding of the epicardial contour into the right ventricular blood pool will incorporate voxels with rapid signal enhancement (due to fast flowing blood in the right ventricle) into the myocardial ROI incurring severe changes in estimated MBF. At stress the genuinely higher MBF obscures the effect of the right ventricular voxels and a significant effect is not observed. These results suggest that a safety tolerance of $MD = 0.3MW$ would be acceptable for segmental analysis of MBF in the healthy myocardium, excluding anteroseptal segment of the resting epicardium, which requires an accuracy of $MD = 0.1MW$

4.3.3. Endo and epicardium Considering the endo and epicardium as separate regions the spread of MBF error increases more rapidly again with increasing MD due to the further decreasing ROI size. A tolerance of $MD \leq 0.2MW$ is now required to avoid significant MBF errors, excepting the anteroseptal segment which sees significant effects even at $MD = 0.1MW$ in the resting epicardium (F-test $p = 0.01$). Considering the mean endocardium and mean epicardium a tolerance of $MD \leq 0.3MW$ is sufficient to avoid significant MBF errors.

4.4. Inter and intra observer variability

We required our contour errors to be random, in the absence of knowledge of any more systematic form of error population, whilst maintaining a smooth circular form, in that one would not expect a manual contourer to deliberately generate sharp corners or high frequency oscillations in the contour. To assess whether our simulated contour errors were representative of the errors that manual contourers make we compared our error distributions with inter- and intra-observer contour error distributions. Figures 8a and 8b show that the distributions of contour errors that we simulated are similar to inter- and intra-observer

variabilities, with the best visual match being between the $0.4MW$ & $0.5MW$ MD simulations for inter-observer variability and between the $0.2MW$ and $0.3MW$ MD simulations for intra-observer variability. Kolmogorov-Smirnoff tests confirm this observation showing that there is insufficient evidence to reject the null hypothesis that the inter/intra-observer contour errors were drawn from the same underlying distribution as the simulated contour errors for all maximum deviation values except the $MD = 0.1MW$ case. Analysis of the separate stress, rest, endo- and epi-contours did not yield evidence to alter this conclusion.

It is evident from figure 8 that the inter-observer variation in contour errors was much broader than the intra-observer variation. This is an expected result. The bimodal shape of the inter-observer contour error distribution is due to one of the observers being consistently more conservative in their contouring. This was evident when the endo- and epi-contour cases were considered separately, yielding normal distributions of contour errors with positive means for the epi-contour errors and negative means for the endo-contour errors.

4.5. Units of Contour Error

The data in this study is expressed in terms of the maximum deviation (MD) that a contour is allowed to deviate from its 'true' value. This is the most meaningful measure of contour error available as we have shown that recognized segmentation metrics do not correlate with MBF error and it encapsulates clearly the concept behind our philosophy of simulating contour errors and can be directly interpreted as a maximum allowable contour error limit for any contouring algorithm. Expressing MD in terms of voxels would be inadequate because voxel size changes from MRI acquisition to acquisition. But converting results into absolute spatial measures (*mm* or mm^2) is not appropriate either as any cohort of cardiac MR images will contain a range of heart sizes, thus a given absolute contour error will have a more profound effect on a smaller heart than on a large one. Therefore our measures of contour error are expressed as a fraction of the mean myocardial width (MW) of the specific heart being considered. Expressing contours this way ensures that a given contour error has the same effect across the entire dataset in terms of its relative geometric change to the myocardial contour with respect to the myocardium.

4.6. Limitations

This study has been carried out on healthy volunteers only. The inclusion of ischaemic patients, whose MBF is compromised, would incorporate a confounding factor into the data. We have recommended contour error limits under which healthy MBF estimates do not significantly vary. The effect of such errors on ischaemic patients has not been investigated and the tolerances required there may differ.

MBF varies between systole and diastole (Radjenovic et al. 2010). To exclude this complicating factor from the study only systolic images were analyzed. The systolic myocardium is thicker and thus easier to contour thus providing a more trustworthy contour reference standard. The contour accuracy required for analyzing diastolic data, where the myocardial wall is thinner, may be higher thus our conclusion should strictly only be applied

to systolic data. However, our contour errors are expressed as a fraction of the mean myocardial width (MW) and therefore might feasibly be applied to diastolic data if the distribution of myocardial widths around the myocardial circumference is similar in diastole and systole.

Our manual contouring only allowed for rigid translations in the motion-correction, which is consistent with clinical practice at our institution. The inherent assumption is that there is no rotational or in-plane motion of the heart, which may not be true during breathing motion. However, as breathing motion has been minimized over the first-pass by the adopted breath-holding strategy the errors induced due to this assumption should be minimal. The alternative approach of manually contouring each image in the dynamic series is much more time consuming and is difficult in low contrast images, where the myocardium and surrounding tissues can be indistinguishable. There is no reason therefore to suppose that our method incurs worse errors than a method incorporating an independent contour for each time step.

We have simulated contour errors by allowing random variations evenly around the contour and we have shown that these simulations generate a similar distribution of contour errors to inter- and intra-observer variability distributions. It may be the case that contour errors are more likely over certain regions (e.g. where there are more poorly defined edges) than others thus an even distribution of random contours is not the best simulation. However, in this case our simulations will overestimate the MBF errors due to a given MD and so our suggested tolerance levels can be treated as a conservative limit which may yield lower, but not higher, errors in MBF than reported here.

Our implementation of the method proposed by (Larsson et al. 1996) to convert signal intensities to concentrations has two important limitations. Firstly this method was originally validated for an inversion recovery sequence (Fritz-Hansen et al. 1996), (Fritz-Hansen et al. 2008). Although the adaptation of the method to a saturation recovery sequence is mathematically simple the method has not been separately validated for this sequence. Secondly the use of an assumed T_1 for blood may introduce errors in to the MBF estimation process. To verify the robustness of this approach to errors in the assumed T_1 value we estimated MBF values from a subset of 10 volunteers whose signal intensity vs. time curves had been converted to concentrations using a range of assumed T_1 values. The assumed T_1 values were chosen to cover the 95% confidence interval of a weighted average of native cardiac blood T_1 measured by (Flacke et al. 2001), (Klein et al. 2004), (Messroghli et al. 2004), (Sharma et al. 2006). Over this range of assumed T_1s median MBFs over all patients did not fall outside of the interquartile range of MBFs corrected with the 'true' assumed blood T_1. The spread of MBF values estimated with this 'true' MBF was comparable with the literature (Pack et al. 2008), (Fritz-Hansen et al. 2008). This result showed that errors in MBF due to assumed blood T_1 are less pronounced than the inherent inter-patient variation in MBF thus MBF estimates are relatively robust to errors in assumed blood T_1 (Biglands et al. 2010).

5. Conclusion

Myocardial contour errors have been simulated for estimation of MBF. The relationships between segmentation error and MBF error has been described and explained in terms of cardiac physiology in healthy volunteers. The segmentation evaluation metrics considered are not correlated with random MBF errors thus neither measure fully evaluates whether the contours are fit for purpose. Thus contouring algorithms for DCE-MRI myocardial perfusion should ideally be evaluated by both geometric segmentation metrics and in terms of MBF error. Comparison of segmentation algorithms evaluated with segmentation evaluation metrics with those evaluated with MBF error is not possible unless the segmentation errors are systematic in nature.

For healthy volunteers imaged at systole with the scan parameters used in this study a significant deviation from the range of true MBF values can be avoided if contour errors are randomly distributed with a maximum deviation (MD) of $0.4MW$, when considering the whole myocardium, $0.3MW$, when considering 6 radial segments as prescribed by the AHA (Cerqueira et al. 2002), and $0.2MW$ if the myocardium is further subdivided into endo and epicardium. The exception is the anteroseptal region which required $\leq 0.1MW$ for the AHA model and $< 0.1MW$ for endo and epicardial segments. In general more significant MBF errors were observed when the myocardial contours passed outside, rather than inside, the myocardium suggesting that conservative contour drawing (remaining within the myocardium) will generate smaller errors than generous contouring. However it should be noted that our evaluation was carried out on healthy volunteers only, and conservative contouring could exclude important sub-myocardial defects in ischaemic patients.

Acknowledgments

This work was supported by the British Heart Foundation Programme Grant (RG/05/004). This report/article presents independent research commissioned by the National Institute for Health Research (NIHR). The views expressed in this publication are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health. The authors would like to acknowledge David Buckley and Steven Sourbron for their insightful comments on the perfusion quantitation method.

6. References

Adluru G, DiBella E & Whitaker R 2006 *in* 'Biomedical Imaging: Nano to Macro, 2006. 3rd IEEE International Symposium on' IEEE pp. 133–136.

Ball R & Bache R 1975 *J. Clin. Invest.* **55**(1), 43–49.

Beauchemin K 1998 *Can. J. Rem. Sens.* **24**(1), 3–8.

Biglands J, Larghat A, Plein S, Buckley D, Jerosch-Herold M, Magee D, Boyle R & Radjenovic A 2010 *in* 'Proceedings of the 18th International Society of Magnetic Resoannce In Medicine Scientific Meeting, ISMRM 2010, Stockholm, Sweden.'.

Brent R 1973 *Algorithms for Minimization Without Derivatives* Prentice-Hall.

Cerqueira M D, Weissman N J, Dilsizian V, Jacobs A K, Kaul S, Laskey W K, Pennell D J, Rumberger J A, Ryan T J & Verani M S 2002 *J. Cardiov. Magn. Reson.* **4**(2), 203–210.

Cheong L H & et al. 2003 *Phys. Med. Biol.* **48**(5), N83.

Christian T F, Rettmann D W, Aletras A H, Liao S L, Taylor J L, Balaban R S & Arai A E 2004 *Radiology* **232**(3), 677–684.

Dice L R 1945 *Ecology* **26**(5), 297–302.

Flacke S J, Fischer S E & Lorenz C H 2001 *Radiology* **218**(3), 703–710.

Fritsch F N & Carlson R E 1980 *SIAM J. Num. Anal.* **17**, 238–246.

Fritz-Hansen T, Hove J D, Kofoed K F, Kelbaek H & Larsson H B W 2008 *J. Magn. Reson. Imag.* **27**(4), 818–824.

Fritz-Hansen T, Rostrup E, Larsson H B W, Sondergaard L, Ring P & Henriksen O 1996 *Magn. Reson. Med.* **36**(2), 225–231.

Jerosch-Herold M, Seethamraju R T, Swingen C M, Wilke N M & Stillman A E 2004 *J. Magn. Reson. Imag.* **19**, 758–770.

Jerosch-Herold M, Wilke N & Stillman A 1998 *Med. Phys.* **25**(1), 73–84.

Jerosch-Herold M, Wilke N, Wang Y, Gong G R, Mansoor A M, Huang H, Gurchumelidze S, Stillman A E & Cardiac M R I G 1999 *Int. J. Cardiac. Imag.* **15**(6), 453–464.

Klein C, Nekolla S G, Balbach T, Schnackenburg B, Nagel E, Fleck E & Schwaiger M 2004 *J. Magn. Reson. Imag.* **20**(4), 588–594.

Larsson H B W, FritzHansen T, Rostrup E, Sondergaard L, Ring P & Henriksen O 1996 *Magn. Reson. Med.* **35**(5), 716–726.

Lee D C, Simonetti O P, Harris K R, Holly T A, Judd R M, Wu E & Klocke F J 2004 *Circulation* **110**(1), 58–65.

Massey F 1951 *J. Am. Stat. Assoc.* **46**(253), 68–78.

Messroghli D R, Radjenovic A, Kozerke S, Higgins D M, Sivananthan M U & Ridgway J P 2004 *Magn. Res. Med.* **52**(1), 141–146.

Pack N A, DiBella E V R, Rust T C, Kadrmas D J, McGann C J, Butterfield R, Christian P E & Hoffman J M 2008 *J. Cardiov. Magn. Reson.* **10**, 15.

Parks C M & Manohar M 1983 *J. Appl. Physiol.* **54**(6), 1641–1652.

Radjenovic A, Biglands J D, Larghat A L, Ridgway J, Ball S G, Greenwood J P, Jerosch-Herold M & Plein S 2010 *Magn. Reson. Med.* .

Rohrer M, Bauer H, Mintorovitch J, Requardt M & Weinmann H J 2005 *Invest. Radiol.* **40**(11), 715–24.

Santarelli M, Positano V, Michelassi C, Lombardi M & Landini L 2003 *Med. Eng. Phys.* **25**, 149–159.

Sharma P, Socolow J, Patel S, Pettigrew R I & Oshinski J N 2006 *J. Magn. Reson. Imag.* **23**(3), 323–330.

Stegmann M, Ólafsdóttir H & Larsson H 2005 *Med. Imag. Anal.* **9**, 394–410.

Wagner A, Mahroldt H & Holly T 2003 *Lancet* **361**, 374–379.

Yasnoff W, Mui K & Bacus J 1977 *Pattern. Recogn. Lett.* **9**(4), 217–231.