Investigation into diagnostic accuracy of common strategies for automated perfusion motion correction

Constantine Zakkaroff
John D. Biglands
John P. Greenwood
Sven Plein
Roger D. Boyle
Aleksandra Radjenovic
Derek R. Magee

SPIE.

Investigation into diagnostic accuracy of common strategies for automated perfusion motion correction

Constantine Zakkaroff,[a,*] John D. Biglands,[b,c] John P. Greenwood,[b,c] Sven Plein,[b,c] Roger D. Boyle,[d] Aleksandra Radjenovic,[e] and Derek R. Magee[a]

[a]University of Leeds, School of Computing, Leeds LS2 9JT, United Kingdom
[b]University of Leeds, Division of Medical Physics, Leeds LS2 9JT, United Kingdom
[c]University of Leeds, Multidisciplinary Cardiovascular Research Centre and Leeds Institute of Cardiovascular and Metabolic Medicine, Leeds LS2 9JT, United Kingdom
[d]University of Aberystwyth, Institute of Biological, Environmental and Rural Sciences, Aberystwyth, Ceredigion SY23 3DA, United Kingdom
[e]University of Glasgow, Institute of Cardiovascular and Medical Sciences, Glasgow Cardiovascular Centre, Glasgow G12 8QQ, United Kingdom

Abstract. Respiratory motion is a significant obstacle to the use of quantitative perfusion in clinical practice. Increasingly complex motion correction algorithms are being developed to correct for respiratory motion. However, the impact of these improvements on the final diagnosis of ischemic heart disease has not been evaluated. The aim of this study was to compare the performance of four automated correction methods in terms of their impact on diagnostic accuracy. Three strategies for motion correction were used: (1) independent translation correction for all slices, (2) translation correction for the basal slice with transform propagation to the remaining two slices assuming identical motion in the remaining slices, and (3) rigid correction (translation and rotation) for the basal slice. There were no significant differences in diagnostic accuracy between the manual and automatic motion-corrected datasets ($p = 0.88$). The area under the curve values for manual motion correction and automatic motion correction were 0.93 and 0.92, respectively. All of the automated motion correction methods achieved a comparable diagnostic accuracy to manual correction. This suggests that the simplest automated motion correction method (method 2 with translation transform for basal location and transform propagation to the remaining slices) is a sufficiently complex motion correction method for use in quantitative myocardial perfusion. © *The Authors. Published by SPIE under a Creative Commons Attribution 3.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI.* [DOI: 10.1117/1.JMI.3.2.024002]

Keywords: automated perfusion motion correction; perfusion registration; quantitative perfusion analysis.

Paper 15228RR received Nov. 20, 2015; accepted for publication Apr. 26, 2016; published online May 13, 2016.

1 Introduction

Magnetic resonance imaging (MRI) dynamic contrast enhanced (DCE) myocardial perfusion data have been shown to have a high sensitivity and specificity for diagnosing myocardial ischemia.[1,2] DCE-MRI perfusion datasets can be analyzed to provide quantitative myocardial perfusion estimates, which have also been shown to perform well diagnostically.[3,4] Quantitation of myocardial perfusion requires regions of interest (ROIs) to be drawn over the myocardium and blood pool in every frame in the DCE-MRI dataset to obtain signal enhancement versus time curves. However, in the presence of myocardial displacement due to respiration, this task is prohibitively time consuming for clinical use. The main challenge of manual motion correction is the respiratory motion, causing the imaging plane to pass through an entirely different location (above or below the intended slice location), which breaks the underlying assumption of perfusion analysis. Motion correction methods attempt to remove the respiratory motion from the dataset so that ROIs are required on a single frame only. However, automated motion correction is challenging in DCE-MRI due to poor signal-to-noise ratio in the images, the temporally changing image contrast, and through-plane motion.

The work of McLeish et al.[5] reports that the magnitude of the respiration-induced through-plane motion of the heart is largest in the cranio-caudal direction with smaller displacements in the left–right and anterior–posterior directions; the authors report respiration-induced through-plane translation of up to 23 mm combined with rotation and nonrigid deformation. In the work of Milles et al.,[6] it is pointed out that through-plane motion violates the underlying principle of perfusion imaging, which relies on the idea of tracking the concentration of the contrast agent for the same location in the myocardium over time. In the literature on perfusion series registration, the respiration-induced motion is often mistaken for in-plane two-dimensional (2-D) motion because the shape of the myocardium does not vary significantly around the medial slice, which may create an impression of the respiration motion occurring only in 2-D without through-plane motion. In practice, small magnitudes of through-plane motion can be ignored, which cannot be said about severe cases of through-plane motion. The problem of through-plane motion cannot be solved perfectly with postacquisition methods only, and it has the potential of reducing the sensitivity of perfusion analysis. All postacquisition solutions for automated motion correction in perfusion series aim to offer a sensible compromise to achieve the best estimate for the perfusion-related parameters. Various registration methods based on free-form deformations and collectively known as nonrigid or deformable methods have been reported in the literature over the past years. However, keeping the problem of

---

*Address all correspondence to: Constantine Zakkaroff, E-mail: c.zakkaroff@leeds.ac.uk

through-plane motion in mind, it can be concluded that deformable motion correction is unlikely to improve the diagnostic accuracy of perfusion analysis because it solves an incorrectly posed registration problem. For these reasons, pragmatic simplifications may produce more useful motion correction solutions than complex deformable approaches.

Typically, the accuracy of motion correction algorithms is assessed on the basis of geometric metrics such as the Dice similarity coefficient[7,8] and Hausdorff distance.[9] The Hausdorff distance measures the distance between two subsets of a metric space; Hausdorff distance between two sets is small if every point of either set is close to some point of the other set. In the context of image analysis, the Hausdorff distance metric is used for determining the degree of similarity between two objects when they are superimposed on one another. These evaluation methods require ground-truth motion correction parameters obtained manually against which to compare the automated results. Inevitably, manual definition of ground-truth motion correction is subject to human operator error, and in some cases may be less accurate than automated motion correction. Although automated perfusion motion correction techniques offer tangible benefits, the effects of automated solutions on the clinical outcome measures are rarely assessed.

The work presented here was a retrospective substudy using data from the CE-MARC trial.[1] The CE-MARC trial evaluated the performance of a multiparametric cardiac MR protocol for diagnosing significant coronary artery disease and compared this to single photon-computed emission computed tomography (SPECT) using the reference standard of invasive x-ray angiography. The dataset allowed an evaluation of quantitative MR perfusion estimates diagnostically. The CE-MARC trial[1] dataset provides an ideal opportunity to test motion correction for quantitative perfusion in terms of the diagnosis of myocardial ischemia, which is ultimately the purpose of motion correction in cardiac perfusion DCE-MRI. The aim of our study was to use the clinical gold standard, provided by CE-MARC, to evaluate diagnostically and compare the outcomes of four strategies for motion correction of varying complexity. The strategies included combinations of translation, rotation, and deformable registration approaches as described in the Sec. 2. Our hypotheses were first, that motion correction of the basal slice, with transform propagation to the other slices, provides results that are at least as good as independent motion corrections for each slice; and second, that image translation alone is sufficient to account for respiratory motion for the purposes of quantitative perfusion. We would like to point out that the comparison of registration strategies was made on the basis of their impact on the diagnostic accuracy. Evaluation in terms of diagnostic accuracy provides evidence based on which pragmatic decisions can be made about the necessary complexity and accuracy of motion correction for cardiac perfusion DCE-MRI datasets.

2 Methods

Three short axis slices (basal, medial, and apical) were acquired during the first pass of the contrast agent through the myocardium. The correspondence of the cardiac phase within each spatial location was achieved through electrocardiogram (ECG) gating. Adenosine-induced stress imaging was performed 15 min before rest imaging. Myocardial perfusion MR imaging was performed using bolus intravenous injection of 0.05 mmol/kg of body weight gadopentetate dimeglumine (Magnevist; Schering, West Sussex, England). Images were acquired using a T1-weighted saturation recovery turbo field-echo imaging sequence. A shared (nonslice-selective) saturation pulse was used, giving prepulse delay times of 126 ms, 272 ms, and 418 ms for the basal, middle, and apical slices, respectively. The pulse sequence parameters were as follows: repetition time/echo time, 2.7/1.0 ms; flip angle, 15 deg; sensitivity encoding factor, 2; matrix, 144 3 144; field of view, 320 to 460 mm; slice thickness, 10 mm; and partial Fourier factor, 0.67.

The independent Clinical Trials Unit at the University of Leeds selected 50 patients as a subset of the CE-MARC trial population representative for risk factors and disease severity. All CE-MARC patients underwent x-ray angiography, SPECT, and a multiparametric MR examination. In order to have a consistent dataset, patients with any x-ray angiography defined stenosis severity between 50% and 70% or whose SPECT and x-ray angiography results were discordant were excluded from the study. Patients were recruited from Leeds Teaching Hospitals NHS Trust, Leeds, United Kingdom, and Pinderfields Hospital, Wakefield, United Kingdom. All SPECT and CMR scans were undertaken at Leeds General Infirmary. The study was performed in accordance with the Declaration of Helsinki (October 2000), with all patients providing informed written consent. The study protocol and other relevant documentation were approved by the Leeds (West) Research Ethics Committee.

2.1 Registration Strategies Overview

The registration methods for motion correction in perfusion series compared here draw on the basic frame-by-frame registration approach reported by Li and Sun.[10] In this approach, the motion correction process is started with the first two frames, labeled as fixed and moving images accordingly. For each consecutive registration step, the moving image from the previous step is resampled and treated as the new fixed image. The key novelty of the solution presented here is in the use of average frames as fixed images during registration to avoid the issue of the accumulated registration error, which is common for frame-by-frame registration methods. As noted earlier, with these methods, for each pair of adjacent images, the moving (corrected) image becomes the reference image for the next registration step. Under this scheme, individual registration errors in early frames tend to be exacerbated for the later frames in the series. Use of average images draws on the ideas of groupwise registration published by Cootes et al.[11] The rationale behind the use of average images is that they contain all features from the series, which makes them suitable to be used as the fixed images for the correction of slices from precontrast, right-ventricular contrast, left-ventricular contrast, and contrast wash-out phases of the series. Average images are used as the fixed images in the rigid and deformable versions of motion correction methods. An example of average images and their progressive refinement is given in Fig. 1.

Motion correction was performed on the 50 rest and 50 stress perfusion datasets in accordance with four registration strategies. These strategies vary among themselves in the following aspects:

1. Complexity of the geometric transform: translation-based, rigid, and deformable transforms were used in the experiments. It is common in image registration practice to start motion correction with the simplest possible transform to compensate for gross displacement:
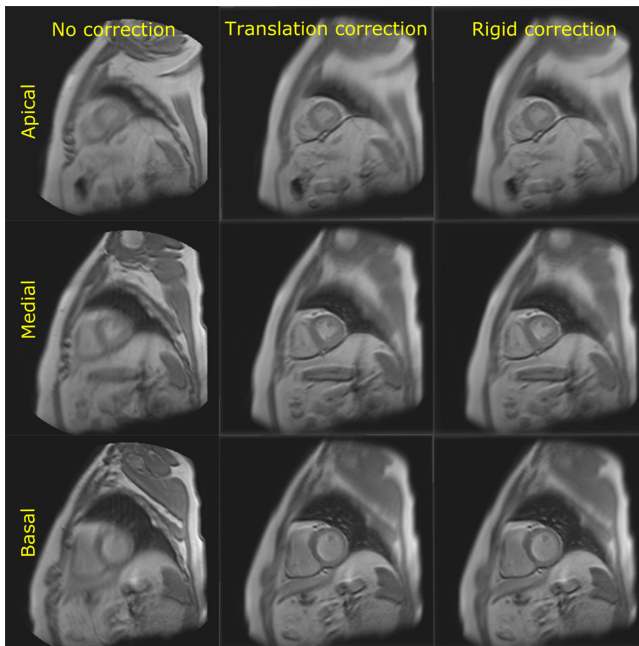
Fig. 1 Examples of average images. Evolution of average images for a stress perfusion series; rows: apical, medial, and basal locations; columns: prior to correction, after translation, and after rigid correction; the improvement of the clarity of the cardiac features from non-corrected average images to translation-corrected images supports the general observation that the bulk of the motion is compensated by translation correction; the rigidly corrected average images also usually show improvement in clarity that may vary from patient to patient.

translation transform. If further correction is required to account for the rotational motion component, the output of the translation correction method is used to initialize the rigid (combination of translational and rotational displacements) correction method. Similarly, the deformable motion correction method needs to be initialized with gross motion correction obtained either with translation or rigid transform. Complex registration solutions in practice form registration pipelines, where motion correction results for distinct stages of the pipeline are used to initialize subsequent motion correction stages. Depending on the nature of the experiments, the results from distinct stages of motion correction can be treated as independent results for distinct types of registration strategies: translation, rigid, and deformable. This was the case with the experiments described here.

2. Choice between the two options of dealing with motion in different slices: one option involves independent motion correction in all slices, basal, medial, and apical, while the other option involves motion correction in the basal slice with transform propagation to the medial and apical slices. It is hypothesized here that it is sufficient to correct the motion only for the basal slice in the whole series, and propagate the transforms to the medial and apical slices. This hypothesis is based on the following observations:

a. Basal slices contain the greatest information density, while the clarity and scale of image features usually progressively decrease from basal to apical slices with the decreasing size of the ROIs. Transform optimization based on any given image similarity metric provides better results for larger objects even with significant initial displacements because the image similarity metric produces stronger gradients for larger features. Thus it is likely that motion correction of basal slices will provide the most accurate results.

b. The perfusion imaging protocol ensures that all three slices in each dynamic frame, basal, medial, and apical, are acquired within 150 ms of each other in the same cardiac cycle. Since respiratory motion is relatively slow, the difference of respiratory motion for the three slices in each frame can be considered negligible. The contractile myocardial motion is much faster than respiratory motion, but with ECG gating functioning correctly, the phase remains consistent for all slices for a given location. This ensures the consistency of spatial position and shape of myocardium with the given location. Thus, transform parameters for all three slices are expected to be similar even in the presence of respiratory motion.

The hypothesis about basal-only slice correction can be applied to translation and rigid motion correction methods; the deformable correction naturally has to involve the optimization of distinct transforms, one for each of the apical, medial, and basal slices within each dynamic frame.

2.2 Common Registration Components

Apart from transform type, its initialization, and parameter propagation, the rest of the registration components were identical for all compared registration methods. Motion correction was achieved through two-level multiresolution registration (half resolution for the first level and full resolution for the second) with a basic gradient-descent optimizer.[12] Nonlinear intensity variations during contrast uptake captured in perfusion series and the intensity values averaged across the whole series call for the use of the mutual information (MI) image similarity metric; Mattes' implementation of the MI image similarity metric[13] was used in all compared registration strategies. Manually defined ROIs were used during registration to exclude irrelevant anatomical features in the images. The ROIs were defined in the key frame of each perfusion series. The key frame is identified visually by the maximal contrast agent concentration in the left ventricle (LV), where image contrast between the blood pool, myocardium, and surrounding structure are optimal. Rectangular ROIs were drawn around the heart, including both LV and right ventricle (RV; as shown in Fig. 2), because the features of the RV are important for recovering the rotational motion. All instances of registration described here were carried out with elliptical masks fitted inside the ROIs. Following are descriptions of the registration strategies used in this research.

2.2.1 Strategy 1: Translation correction for all slices

Pairwise registration of individual basal slices to the maximal contrast basal slice with translation transform was followed by the same procedure for the medial and apical slices and the corresponding medial and apical maximal contrast slices.
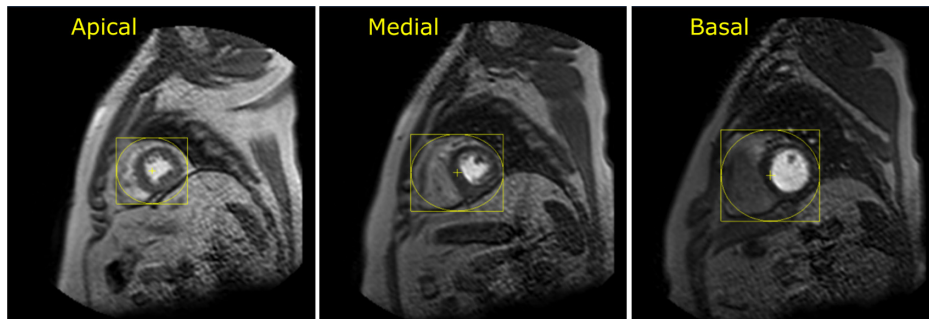
Fig. 2 Examples of manually defined ROIs. An illustration of ROIs in stress perfusion series; the ROIs were defined in a way that aimed to exclude irrelevant features, such as liver and lung tissue, which may move independently of the myocardium; the fixed image masks for registration were obtained from ellipses fitted inside the ROIs; the ROIs and masks were intended to exclude the features irrelevant (or possibly detrimental) to registration.

Pairwise registration starts with the slices adjacent to the maximal contrast frame and continues toward the start and end of the time series as shown in Fig. 3. For each pairwise registration, the initialization of the transform was achieved in the following manner: for the slices on either side of the maximum contrast slice k (positions $k - 1$ and $k + 1$, respectively), the transform was initialized to identity. As the positions of the registered slices, $k - n$ and $k + n$ (where n is the distance between the given slice and the maximal contrast frame) shift from the maximal contrast slice toward the start and end of the series, respectively, the transforms are initialized with the parameters which were obtained from the previously registered slices.

2.2.2 Strategy 2: Translation correction for basal slice with transform propagation

Pairwise registration of individual basal slices to the maximal contrast basal slice with translation transform was followed by the application of the recovered transform for a given basal slice to the lower slices in the corresponding dynamic frames. Transform initialization followed the same pattern as in strategy 1.

2.2.3 Strategy 3: Rigid correction for basal slice with transform propagation

The average image from all basal slices was computed after the application of transforms obtained with strategy 2. Pairwise registration of individual basal slices to the average image of the basal slice with a rigid transform (as shown in Fig. 3) was followed by the application of recovered transforms for a given basal slice to the lower slices in the corresponding dynamic frames. Transform initialization followed the same pattern as in the earlier strategies. The center of rotation was initialized to the center of the LV ROI. The registration algorithm for rigid transformation optimizes the center of rotation parameters along with the translation
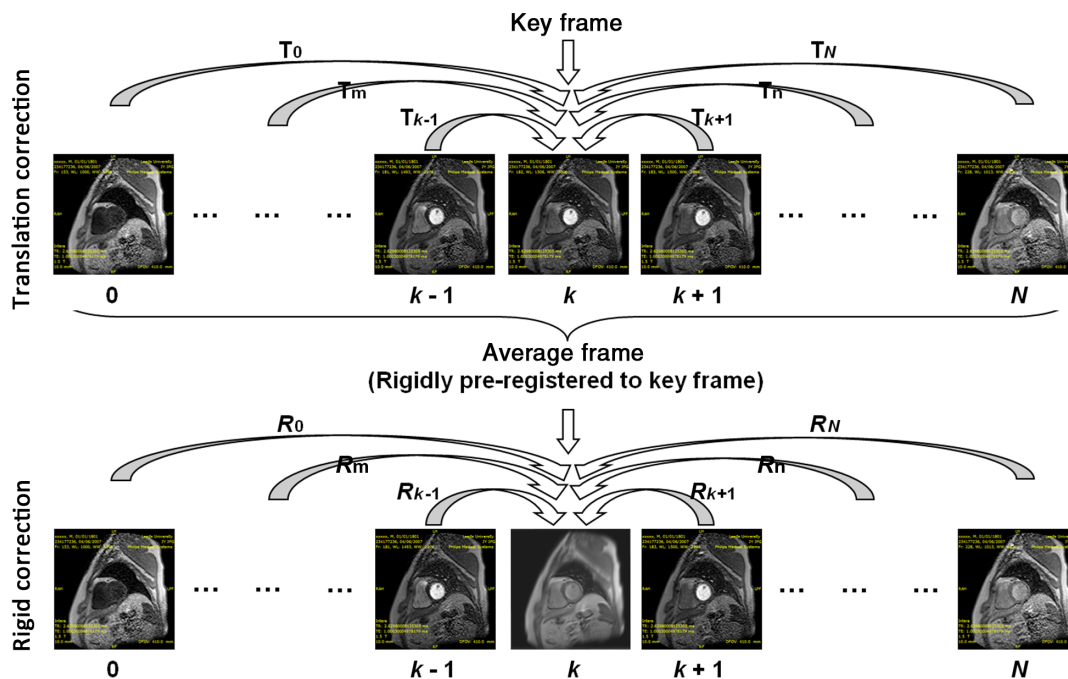

Fig. 3 Perfusion series registration. Perfusion motion correction for basal slice: in the translation correction step, the slices are registered to the basal slice from the key frame; after translation correction, average image is computed and preregistered rigidly to the key frame; during the rigid correction step, the average is used as the fixed image.

parameters. Thus the potentially suboptimal initial center of rotation parameters were adjusted by the transform optimizer to obtain optimal center of rotation position.

2.2.4 Strategy 4: Deformable correction for all slices

The average image from all basal slices was computed after the application of transforms obtained with strategy 3. It is a common requirement for successful registration to supply best-guess initial parameters for any type of registration. As described by Ng and Ibáñez,[14] the cascaded transform initialization in a registration pipeline was used in this study for initializing rigid and deformable transforms. Pairwise registration of individual basal slices to the average image of the basal slice with a B-spline transform[15] was carried out in a manner identical to previously listed strategies. Next, deformable correction was independently carried out for medial and apical slices. Transform initialization followed the same pattern as in the earlier strategies.

2.3 Protocol for Manual Contouring of Perfusion Images

Manually defined contours were used as the reference standard in all perfusion motion correction evaluation experiments. The contours were defined in the dedicated cardiac image analysis software QMass 7.0 by Medis Medical Imaging Systems, Leiden, The Netherlands. For all 50 datasets, LV endocardial, epicardial, and blood pool contours were defined for all frames and slices in rest and stress series. For each perfusion series, a key frame was chosen as the frame exhibiting the maximum contrast between the ventricular blood pool and the surrounding myocardium. First, all three types of contours were defined in the key frame as shown in Fig. 4 (top row); next, the contours from the key frame were copied to the rest of the series; finally, the contours were manually translated within each slice to compensate for respiratory motion. In many images, the observed

Fig. 4 Examples of manually defined contours. An example of manually defined LV endocardial, epicardial, and blood pool contours in the key frame for basal, medial, and apical slices (top row); a typical example of precontrast slices (from the same dataset as above), which are unsuitable for accurate manual contouring (bottom row).

negligible magnitude of motion did not warrant contour adjustment; in most series, this was the case for the frames acquired during breath-hold. Breath-holding in the CE-MARC study perfusion series was timed to coincide with the first pass of the contrast agent through the myocardium.

The approach to manual motion correction described above possesses an inherent limitation because it is based on the assumption that observed motion in perfusion series includes only translation motion, while in practice, through-plane motion and nonrigid motion can be detected in most perfusion datasets. Theoretically, an alternative contouring approach could involve manual contour delineation in each slice and frame. In addition to the significant increase in time and effort this approach would entail, this would not guarantee more accurate contours for the precontrast stages of perfusion series, where low image contrast renders the boundary between myocardium and ventricular blood pool imperceptible, as shown in Fig. 4 (bottom row).

2.4 Evaluation Methods

The evaluation of the proposed perfusion motion correction was carried out in two parts. The first set of experiments evaluated the accuracy of the perfusion motion correction with a general method based on the Dice coefficient metric.[7,8] The experiments in this set involved only motion correction strategies 1 and 2 because the reference contour set was created on the basis of translation motion correction (comparison of rigid and deformable motion correction to the translation-based ground-truth dataset would be meaningless). After motion correction, the Dice coefficient values were calculated for left-ventricular epicardial contours in all frames from the overlap with the contours in the maximal contrast frames. Epicardial contour results are sufficient here because of the "linked" epicardial and endocardial contour translation in the manual contour dataset.

The second set of experiments compared all four strategies for motion correction versus manual motion correction in the context of quantitative perfusion analysis. Registration for each frame and slice was calculated and stored as a vector deformation field (VDF). Inverse VDFs were calculated for the full dataset, because image registration causes interpolation errors in voxel values. Using the inverse VDFs, the contours at the maximum enhancement reference frame (and slice where relevant) were propagated to the rest of the dataset. The myocardium was subdivided into equidistant circumferential regions corresponding to different coronary arteries, as recommended by the American Heart Association (AHA).[16] The registered contours were then used to generate signal intensity versus time curves for the blood pool and the myocardium. Dual-acquisition and dual-bolus methods are designed to deal with the nonlinear relationship between the measured MR signal intensity and the concentration of contrast agent in the heart. In the reported study, this problem was addressed by converting signal intensities to concentration values using the MRI pulse sequence equation and an assumed native T1 value for blood,[17] as described in the work of Biglands et al.[18] The concentration curves from each AHA segment were then analyzed to generate myocardial blood flow (MBF) estimates using Fermi-constrained deconvolution.[19] Myocardial perfusion reserve (MPR) values were then calculated for each AHA segment by dividing the stress and rest MBF estimates. MPR values calculated after manual motion correction were used as the reference diagnostic values. MPR values obtained after the listed motion correction methods were compared against the reference
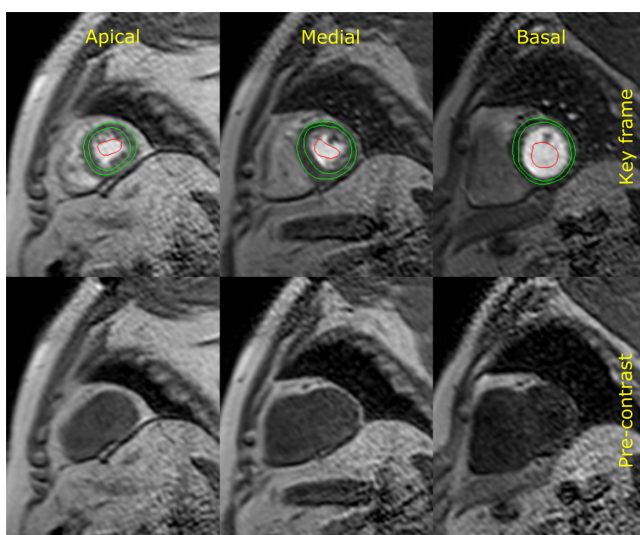
MPR values to judge the performance of a particular motion correction method. Cascaded and pairwise evaluation of registration steps would be meaningless, as it is inappropriate to assume each step represents the reference correction transform. In order to assess diagnostic accuracy, the minimum segmental MPR score was then used for each patient to generate a receiver operator characteristic (ROC) curve using the x-ray/SPECT perfusion consensus diagnosis as the gold standard. Individual ROC curves were generated in this way for each motion correction method. Each curve was then compared with the ROC curve obtained using manual motion correction by means of a DeLong, DeLong, Clarke-Pearson nonparametric comparison of ROC area under the curve (AUC) values.[20]

The automated motion correction strategies were implemented on the basis of the Insight Registration and Segmentation Toolkit libraries.[12] The registration correction experiments can be run on a typical modern workstation in serial mode or in batch mode on a computational node. The registration experiments were performed on the University of Leeds computing node of the White Rose Grid for high performance computing. The translation correction for a single dataset was typically achieved in less than a minute. The deformable registration of all 50 datasets scheduled to run in parallel was typically achieved in less than an hour. The general trend for increased availability of computing power makes the computational expense for a single dataset registration negligible.

3 Results

A visual example of perfusion motion correction is shown in Fig. 5; the images were obtained by extracting reformatted profiles along the XZ image dimension from the 2-D slices for a given spatial location stacked into pseudo-three-dimensional (3-D) images (2-D + time). The extracted profiles were defined to pass through the center of manually defined ROIs shown in Fig. 5 to include the relevant cardiac features and contrast uptake events. The example shows the reformatted profiles for the series before motion correction, after translation correction, and after rigid correction. The main indicator of motion is the interventricular septum. In addition, the images clearly show magnitude of motion with respect to the precontrast, right-ventricular contrast, left-ventricular contrast, and washout phases of the series. In the images prior to correction, the magnitude of respiratory motion can be observed from the changing position of the interventricular septum and ventricles,
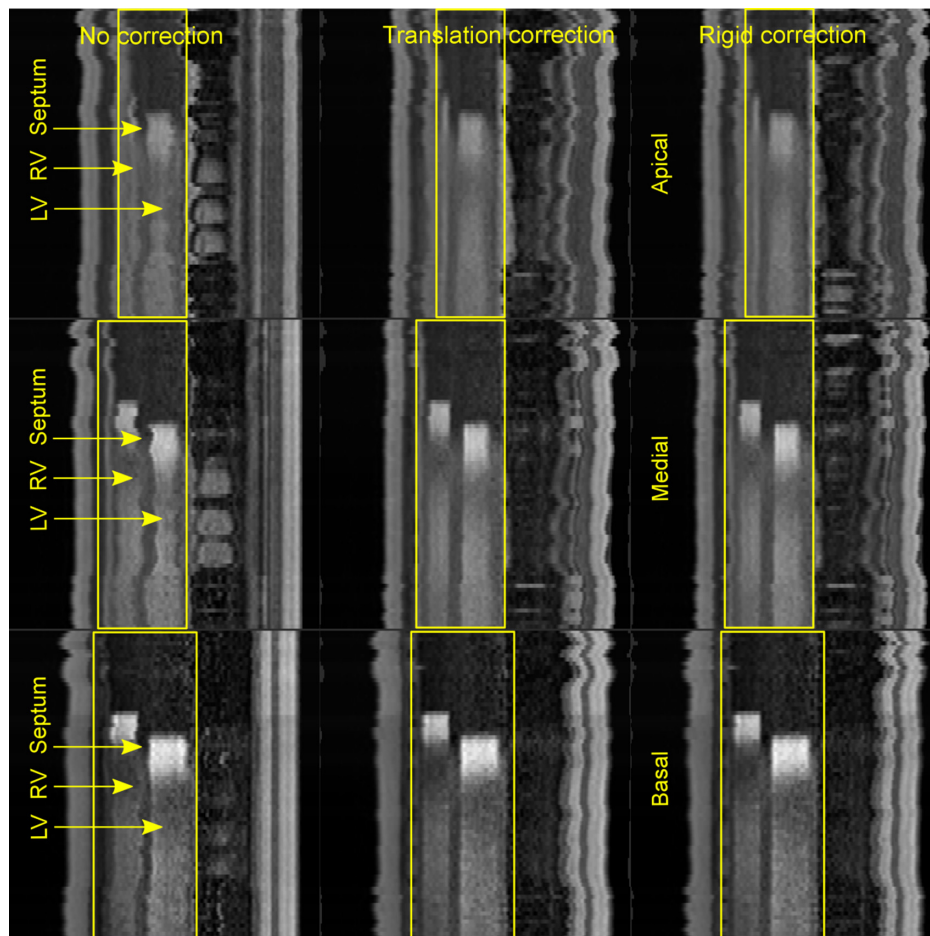

Fig. 5 Visual comparison of perfusion correction stages. Stacked reformatted images for stress perfusion series; rows: apical, medial, and basal locations; columns: before correction, after translation, and after rigid correction; the ROIs are shown with yellow boxes; within each ROI, the RV appears on the left, interventricular septum in the center, and the LV on the right; the last two columns indicate that the bulk of respiratory motion was recovered during translation correction; the interventricular septum is the most obvious feature that can be used for visual validation of motion correction.

while the series after the first correction stage show that the bulk of motion has been recovered with translation transform. The series after the rigid correction stage in this example are very similar to the images after translation correction. This observation is consistent with the examples of average images and their refinement observed after each correction stage is given in Fig. 1, where it is demonstrated how the clarity of the cardiac features in the ROI increases starting from the average images prior to correction, after translation and rigid correction steps. The average images after deformable registration do not show significant improvement in image clarity. The images of the series after deformable correction are not included in this example because visually, they appear identical to the images after rigid correction.

3.1 All-Slice Correction Versus Basal-Only Correction Results

The comparison between the all-slice (strategy 1) and basal-only (strategy 2) variants of perfusion motion correction is based on the Dice metric values (as shown in Table 1) obtained from the comparison of the reference contours and the key frame contours after the application of translation transforms for a given frame and slice.

A two-sample t-test suggests that the difference between the means for the all-slice and basal-only correction strategies ($\mu1$ and $\mu2$, respectively) is not statistically significant for both rest ($\mu1 = 0.957$, $\mu2 = 0.952$, $t = 1.452$, and $p = 0.075$) and stress ($\mu1 = 0.946$, $\mu2 = 0.938$, $t = 1.037$, and $p = 0.151$) series correction. The important finding here is that basal-only correction with transform propagation provides motion correction, which is as accurate as the all-slice correction strategy.

There were no outright failures detected upon visual examination, which was performed in a manner as shown in Fig. 5. The extremely small p-values are explained by the fact that every t-test involved thousands of samples (on the order of 60 to 80 slices for three locations for all 50 patients).

3.2 Diagnostic Evaluation

The ROC curves in Fig. 6 compare the four variants of perfusion motion correction versus manual perfusion correction in terms of the ability to diagnose myocardial ischemia. The AUC values with the associated confidence intervals are shown in Table 2. These values suggest that the results for all four variants of perfusion motion correction are equivalent to manual correction. The cut-off values that achieved the highest sensitivity and specificity for the diagnostic test were as follows: manual correction = 1.34, translation for all slices correction = 1.31, translation for basal slice with transform propagation correction = 1.37, rigid correction = 1.20, and deformable correction = 1.41.

4 Discussion

The study described in this article compares a number of automated perfusion motion correction strategies of various levels of complexity. The proposed solution relies on the use of average images as the fixed images during registration. In the method presented in this article, the average images undergo refinement during the translation-based, rigid, and deformable stages of motion compensation. The MI image similarity metric allows the modeling of complex mappings between pixel intensities, where the actual form of dependency between the intensities in the registered images does not need to be specified. The use of the MI image similarity metric allows the average images to be used as fixed images during correction of perfusion slices from precontrast, RV contrast, LV contrast, and wash-out parts of the perfusion series. The evaluation results prove the validity of the correction approach with translation for basal slices with subsequent transform propagation to medial and apical slices. The Dice metric values for LV endocardial overlap and the ROC curves for MPR scores both confirm the initial hypothesis that the recovery of translation motion in the basal slice is at least as good if not better than translation correction for all slices. The motion correction provides more reliable results with basal images due to the larger scale of the ventricles and other image features, and the transforms from the basal slice can be applied to the medial and apical slices because of the negligible magnitude of respiratory motion observed between the acquisitions of basal, apical, and medial slices within one cardiac cycle.

In most cases, the motion correction based on the rigid 2-D transform provided reliable results. However, during the initial stages of prototyping and exploration, it was observed that the rate of failures was significantly reduced when the translation parameters in the rigid transform were fixed. Most importantly, this study confirms that rigid and deformable correction strategies do not enhance diagnostic accuracy. In practice, the rigid and deformable automated motion correction methods in perfusion analysis are often used incorrectly as a means to deal with myocardium shape variation induced by through-plane motion. Through-plane motion compensation violates the underlying principle of perfusion imaging, which relies on the idea of tracking the concentration of the contrast agent for the same location in the myocardium over time. In such cases, the automated motion compensation methods force the images to show two different slices through the myocardium to be registered to each other. The main message of this study is that the type of automated motion correction does not affect the diagnostic accuracy. In addition, it must be noted that postacquisition image processing methods such as image registration are not suitable for solving the problem of through-plane motion, which breaks the underlying assumptions of quantitative perfusion analysis. Instead, this problem is more suitable for being tackled with methods applicable during image acquisition, such as prospective tissue tracking. It is yet to be explored how motion correction methods can be adapted to emerging 3-D myocardial perfusion techniques such as reported in the work of Jogiya et al.[21] and Manka et al.[22]

The difficulty in evaluating motion correction in DCE-MRI was a key motivation for using clinical diagnosis as a reference standard. Although we are evaluating at the contouring stage and the final diagnosis stage, we are not evaluating at the time versus signal intensity curve stage. If there were a single time point registration error causing "spikes" in these curves, or systematic translational shifts that incorporated blood pool voxels in the myocardial curve, then these could cause errors in MBF. However, it is not clear how a summary signal intensity value could detect these changes. Average values over time would be insensitive to localized changes and would also require normalization for signal gain. Thus, a summary measure would be needed to look at systematic differences in curves between registration methods.

The limitations of this study involved the comparison against translation-only manual motion compensation. However, the ROC curves for MPR scores provide strong evidence of the

Table 1 Dice metric values for strategies 1 and 2.

| Patient # | Rest | | Stress | |
| | Strategy 1 | Strategy 2 | Strategy 1 | Strategy 2 |
	Dice average (st. dev)	Dice average (st. dev)	Dice average (st. dev)	Dice average (st. dev)
1	0.947 (0.055)	0.945 (0.048)	0.966 (0.026)	0.961 (0.034)
2	0.940 (0.050)	0.923 (0.070)	0.956 (0.047)	0.948 (0.047)
3	0.969 (0.023)	0.967 (0.029)	0.957 (0.023)	0.950 (0.031)
4	0.972 (0.020)	0.974 (0.021)	0.969 (0.017)	0.967 (0.025)
5	0.971 (0.023)	0.971 (0.021)	0.948 (0.028)	0.943 (0.039)
6	0.975 (0.018)	0.961 (0.029)	0.963 (0.019)	0.944 (0.041)
7	0.976 (0.019)	0.977 (0.019)	0.779 (0.123)	0.716 (0.112)
8	0.960 (0.028)	0.963 (0.025)	0.934 (0.072)	0.957 (0.031)
9	0.955 (0.033)	0.959 (0.030)	0.938 (0.050)	0.939 (0.042)
10	0.946 (0.029)	0.947 (0.035)	0.958 (0.025)	0.957 (0.022)
11	0.975 (0.019)	0.970 (0.024)	0.884 (0.064)	0.775 (0.090)
12	0.935 (0.051)	0.912 (0.072)	0.892 (0.077)	0.850 (0.090)
13	0.947 (0.039)	0.929 (0.059)	0.927 (0.047)	0.890 (0.073)
14	0.968 (0.017)	0.964 (0.023)	0.958 (0.030)	0.954 (0.032)
15	0.956 (0.032)	0.961 (0.027)	0.927 (0.031)	0.911 (0.038)
16	0.952 (0.022)	0.945 (0.030)	0.951 (0.032)	0.929 (0.041)
17	0.983 (0.012)	0.973 (0.025)	0.955 (0.040)	0.950 (0.043)
18	0.958 (0.028)	0.970 (0.018)	0.962 (0.021)	0.969 (0.022)
19	0.938 (0.059)	0.942 (0.043)	0.942 (0.042)	0.949 (0.041)
20	0.932 (0.041)	0.938 (0.039)	0.968 (0.020)	0.963 (0.022)
21	0.961 (0.032)	0.963 (0.037)	0.951 (0.038)	0.955 (0.033)
22	0.969 (0.018)	0.964 (0.023)	0.964 (0.024)	0.960 (0.026)
23	0.933 (0.072)	0.928 (0.071)	0.928 (0.066)	0.942 (0.045)
24	0.964 (0.026)	0.958 (0.026)	0.954 (0.027)	0.935 (0.043)
25	0.940 (0.055)	0.930 (0.065)	0.934 (0.067)	0.941 (0.061)
26	0.974 (0.019)	0.976 (0.019)	0.971 (0.020)	0.969 (0.027)
27	0.965 (0.023)	0.965 (0.029)	0.944 (0.033)	0.944 (0.035)
28	0.966 (0.044)	0.964 (0.041)	0.961 (0.030)	0.959 (0.037)
29	0.974 (0.019)	0.970 (0.028)	0.965 (0.027)	0.958 (0.037)
30	0.971 (0.021)	0.963 (0.033)	0.961 (0.026)	0.959 (0.031)
31	0.971 (0.028)	0.959 (0.045)	0.974 (0.020)	0.970 (0.020)

Table 1 (*Continued*).

| Patient # | Rest | | Stress | |
| | Strategy 1 | Strategy 2 | Strategy 1 | Strategy 2 |
	Dice average (st. dev)	Dice average (st. dev)	Dice average (st. dev)	Dice average (st. dev)
32	0.945 (0.056)	0.940 (0.077)	0.924 (0.043)	0.901 (0.053)
33	0.961 (0.026)	0.961 (0.027)	0.943 (0.037)	0.944 (0.050)
34	0.940 (0.084)	0.929 (0.102)	0.943 (0.057)	0.940 (0.053)
35	0.951 (0.041)	0.935 (0.052)	0.949 (0.036)	0.945 (0.038)
36	0.978 (0.014)	0.977 (0.017)	0.967 (0.031)	0.966 (0.029)
37	0.969 (0.028)	0.965 (0.033)	0.925 (0.097)	0.947 (0.049)
38	0.965 (0.027)	0.964 (0.028)	0.964 (0.023)	0.959 (0.025)
39	0.969 (0.034)	0.963 (0.034)	0.946 (0.048)	0.934 (0.054)
40	0.934 (0.078)	0.909 (0.084)	0.927 (0.081)	0.932 (0.067)
41	0.944 (0.064)	0.921 (0.084)	0.913 (0.086)	0.908 (0.058)
42	0.954 (0.034)	0.957 (0.029)	0.957 (0.034)	0.955 (0.041)
43	0.954 (0.035)	0.952 (0.033)	0.960 (0.037)	0.965 (0.033)
44	0.972 (0.021)	0.957 (0.031)	0.974 (0.017)	0.961 (0.030)
45	0.961 (0.023)	0.930 (0.050)	0.965 (0.036)	0.957 (0.040)
46	0.936 (0.053)	0.914 (0.075)	0.948 (0.046)	0.943 (0.051)
47	0.954 (0.045)	0.959 (0.037)	0.972 (0.019)	0.969 (0.019)
48	0.911 (0.147)	0.937 (0.071)	0.945 (0.044)	0.933 (0.060)
49	0.970 (0.022)	0.958 (0.024)	0.962 (0.024)	0.953 (0.035)
50	0.958 (0.047)	0.964 (0.041)	0.949 (0.051)	0.945 (0.052)

Average Dice metric values (and standard deviation values) for all contours for a given patient for strategies 1 and 2. Dice metric values were calculated for the reference contours and the key frame contours after the application of translation transforms for a given frame and slice.

overall soundness of the study described here. X-ray angiography is limited as a reference standard for myocardial ischemia because the degree of stenosis is not always linked to a reduction in coronary, or ultimately MBF. This study used a consensus diagnosis between SPECT and x-ray angiography to obtain the best reference standard available with the CE-MARC dataset; however, measurements could have been improved if fractional flow reserve pressure wire measurement had been used.

It is important to note that the diagnostic accuracies achieved in this study are not clinically representative. Ambiguous cases, where x-ray and SPECT measurements did not agree, were removed. This created a reliable reference standard for our comparison study, but our results cannot be considered representative of clinical diagnostic performance.
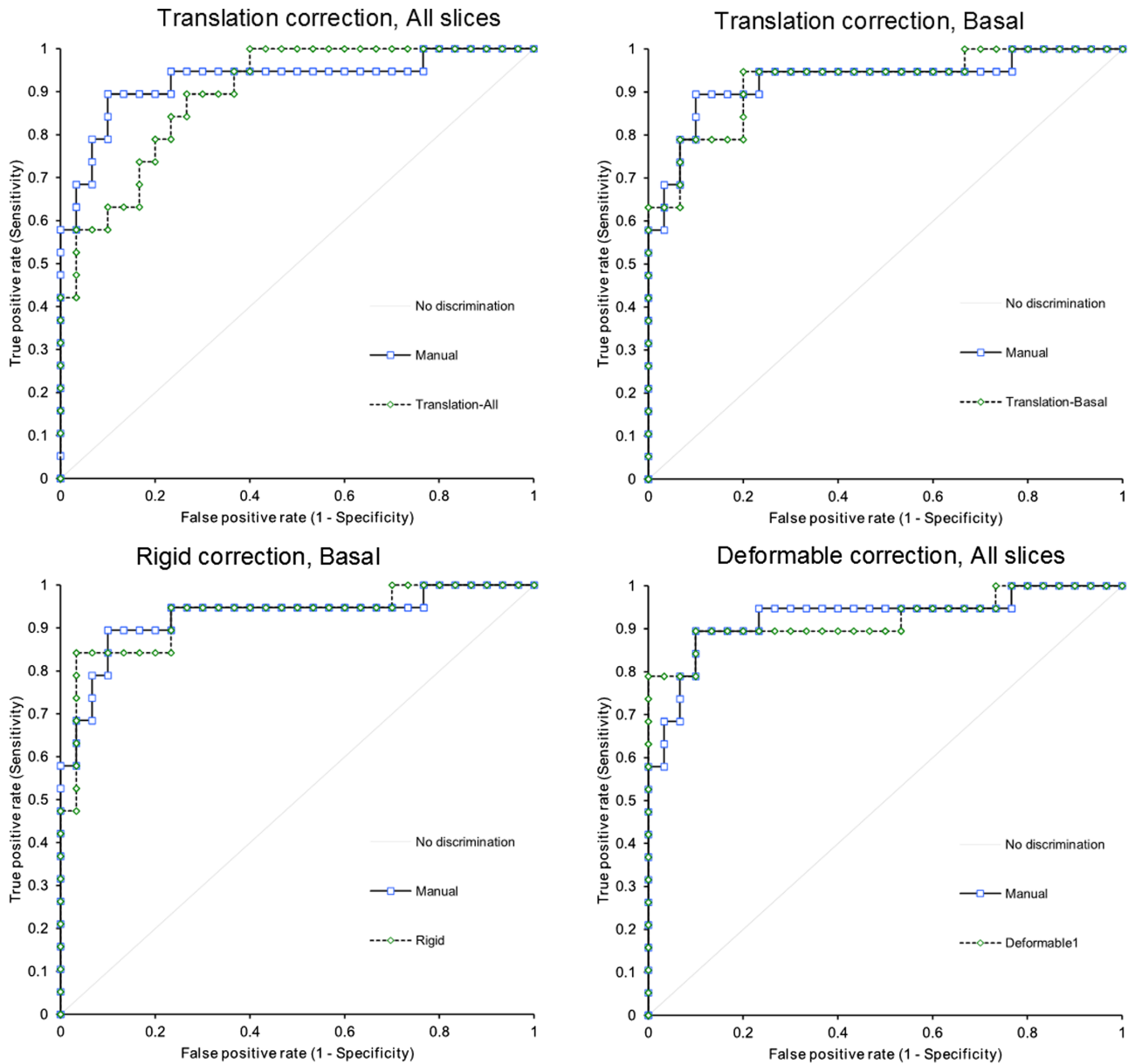
Fig. 6 MPR ROC curves. ROC curves for MPR scores generated for the four variants of perfusion motion correction: translation correction for all slices, translation correction for basal slices, rigid correction for all slices, and deformable correction for all slices; the four variants are compared to the ROC curve generated for the manual motion correction.

Table 2 ROC curve statistics.

Manual	Strategy 1 (translation, all slices)	Strategy 2 (translation, basal slice)	Strategy 3 (rigid, basal slice)	Strategy 4 (deformable, all slices)
AUC 0.93	AUC 0.89	AUC 0.92	AUC 0.93	AUC 0.92
	$p = 0.41$	$p = 0.88$	$p = 1.00$	$p = 0.92$
	(0.81, 0.98)	(0.84, 1.00)	(0.84, 1.00)	(0.83, 1.00)

AUC values and the corresponding confidence intervals for the four perfusion motion correction variants compared to manual motion correction; the AUC values indicate that all four perfusion correction variants are at least as good as manual motion correction. The confidence interval values given in parentheses are the lower and upper 95% confidence interval limits on the AUC values.

5 Conclusions

This study has shown that automated motion correction does not significantly change the diagnostic accuracy of quantitative myocardial perfusion imaging for diagnosing myocardial ischemia. This provides important evidence in support of the use of such methods in practice. Furthermore, this study has shown that the apparently more sophisticated method of deformable registration does not improve diagnostic accuracy above that of simple rigid registration.

References

1. J. P. Greenwood et al., "Cardiovascular magnetic resonance and single-photon emission computed tomography for diagnosis of coronary heart disease (CE-MARC): a prospective trial," *Lancet* **379**(9814), 453–460 (2012).
2. J. Schwitter et al., "MR-IMPACT II: magnetic resonance imaging for myocardial perfusion assessment in coronary artery disease trial: perfusion-cardiac magnetic resonance vs. single-photon emission computed tomography for the detection of coronary artery disease: a comparative," *Eur. Heart J.* **34**(10), 775–781 (2013).
3. A. Huber et al., "Magnetic resonance perfusion of the myocardium: semiquantitative and quantitative evaluation in comparison with coronary angiography and fractional flow reserve," *Invest. Radiol.* **47**(6), 332–338 (2012).
4. A. R. Patel et al., "Assessment of advanced coronary artery disease: advantages of quantitative cardiac magnetic resonance perfusion analysis," *J. Am. Coll. Cardiol.* **56**(7), 561–569 (2010).
5. K. McLeish et al., "A study of the motion and deformation of the heart due to respiration," *IEEE Trans. Med. Imaging* **21**(9), 1142–1150 (2002).
6. J. Milles et al., "Fully automated motion correction in first-pass myocardial perfusion MR image sequences," *IEEE Trans. Med. Imaging* **27**(11), 1611–1621 (2008).
7. L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology* **26**(3), 297–302 (1945).
8. T. Sørensen, "A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons," *Biol. Skr.* **5**(4), 1–34 (1948).
9. D. P. Huttenlocher, G. A. Klanderman, and W. J. Rucklidge, "Comparing images using the Hausdorff distance," *IEEE Trans. Pattern Anal. Mach. Intell.* **15**(9), 850–863 (1993).
10. C. Li and Y. Sun, "Nonrigid registration of myocardial perfusion MRI using pseudo ground truth," *Med. Image Comput. Comput. Assist. Interv.* **12**(Pt 1), 165–172 (2009).
11. T. F. Cootes et al., "Computing accurate correspondences across groups of images," *IEEE Trans. Pattern Anal. Mach. Intell.* **32**(11), 1994–2005 (2010).
12. L. Ibáñez et al., *The ITK Software Guide*, 2nd ed., Kitware, Inc. (2005).
13. D. Mattes et al., "Nonrigid multimodality image registration," *Proc. SPIE* **4322**, 1609 (2001).
14. L. Ng and L. Ibáñez, "Medical image registration: concepts and implementation," in *Insight into Images: Principles and Practice for Segmentation, Registration, and Image Analysis*, T. S. Yoo, Ed., pp. 239–306, A. K. Peters (2004).
15. J. A. Schnabel et al., "A generic framework for non-rigid registration based on non-uniform multi-level free-form deformations," in *Medical Image Computing and Computer Assisted Intervention*, pp. 573–581, Springer-Verlag, Berlin, Heidelberg (2010).
16. M. D. Cerqueira et al., "Standardized myocardial segmentation and nomenclature for tomographic imaging of the heart," *Circulation* **105**(4), 539–542, 2002.
17. T. Fritz-Hansen et al., "Measurement of the arterial concentration of Gd-DTPA using MRI: a step toward quantitative perfusion imaging," *Magn. Reson. Med.* **36**(2), 225–231, 1996.
18. J. D. Biglands et al., "Comparison of the diagnostic performance of four quantitative myocardial perfusion estimation methods used in cardiac MR Imaging: CE-MARC Substudy," *Radiology* **275**(2), 393–402 (2015).
19. M. Jerosch-Herold et al., "Magnetic resonance quantification of the myocardial perfusion reserve with a Fermi function model for constrained deconvolution," *Med. Phys.* **25**(1), 73 (1998).
20. E. R. DeLong, D. M. DeLong, and D. L. Clarke-Pearson, "Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach," *Biometrics* **44**(3), 837–845 (1988).
21. R. Jogiya et al., "Validation of dynamic 3-dimensional whole heart magnetic resonance myocardial perfusion imaging against fractional flow reserve for the detection of significant coronary artery disease," *J. Am. Coll. Cardiol.* **60**(8), 756–765 (2012).
22. R. Manka et al., "Whole-heart dynamic three-dimensional magnetic resonance perfusion imaging for the detection of coronary artery disease defined by fractional flow reserve: determination of volumetric myocardial ischaemic burden and coronary lesion location," *Eur. Heart J.* **33**(16), 2016–2024 (2012).

Constantine Zakkaroff is a lecturer at the Computer Science and Software Engineering. He received his BSc and MSc degrees in computer science from the University of Canterbury, New Zealand, and his PhD from the University of Leeds, United Kingdom. His research interests are in medical image registration, scientific visualisation, large scale physiological modelling, and high performance computing.

John P. Greenwood is a professor of cardiology in the Division of Cardiovascular and Diabetes Research, Leeds University. His main area of research is the use of cardiovascular MR in the assessment of ischaemic heart disease (both chronic stable angina and acute coronary syndromes). This includes the development and validation of comprehensive CMR protocols to assess left ventricular function, myocardial perfusion, tissue oedema, viability, and coronary artery anatomy.

Sven Plein is a professor of cardiology at the University of Leeds, where he leads the Division of Biomedical Imaging. His main research interest is cardiac imaging, in particular cardiac magnetic resonance. The work of his multidisciplinary research group is focused on the development and clinical application of innovative imaging methods in populations at risk of coronary heart disease and heart failure to better detect occult disease and inform patient management.

Roger D. Boyle is a professor of computing in the University of Leeds until 2011, since when he has been imaging specialist at the National Plant Phenomics Centre in Aberystwyth, United Kingdom. His research interests are in model-driven image segmentation.

Derek R. Magee is a lecturer in computer science at the University of Leeds, United Kingdom. He obtained his first class BSc engineering degree from Durham University in 1995, and his PhD from the University of Leeds in 2001. His main areas of interest are image analysis and machine learning in particular applied to medical image and remote sensing data. Additionally, he is chief technical officer of Medical Image Analysis company HeteroGenius Ltd.

Biographies for the other authors are not available.