Published in final edited form as:
NMR Biomed. 2019 June ; 32(6): e4095. doi:10.1002/nbm.4095.

Reproducibility of human cardiac phosphorus magnetic resonance spectroscopy (^{31}P-MRS) at 7T

Jane Ellis[1], **Ladislav Valkovič**[1,2], **Lucian A. B. Purvis**[1], **William T. Clarke**[1,3], and **Christopher T. Rodgers**[1,4]

[1]OCMR, RDM Cardiovascular Medicine, University of Oxford, UK

[2]Institute of Measurement Science, Slovak Academy of Sciences, Bratislava, SVK

[3]Wellcome Centre for Integrative Neuroimaging, FMRIB, Nuffield Department of Clinical Neurosciences, University of Oxford, UK

[4]Wolfson Brain Imaging Centre, Department of Clinical Neurosciences, University of Cambridge, UK

Abstract

Purpose—We test the reproducibility of human cardiac phosphorus magnetic resonance spectroscopy (^{31}P-MRS) at ultra-high field strength (7T) for the first time. The primary motivation of this work was to assess the reproducibility of a 'rapid' 6½ minute ^{31}P three-dimensional chemical shift imaging (3D-CSI) sequence which, if sufficiently reproducible, would allow the study of stress-response processes. We compare this to an established 28 minute protocol, designed to record high quality spectra in a clinically-feasible scan time. Finally, we use this opportunity to compare the effect of per-subject B_0 shimming on data quality and reproducibility in the 6 ½ min protocol.

Methods—10 healthy subjects were scanned on two occasions: one to test the 28 minute 3D-CSI protocol, and one to test the 6½ minute protocol. Spectra were fitted using the OXSA Matlab toolbox. The phosphocreatine to adenosine triphosphate concentration ratio ("PCr/ATP") from each scan was analysed for intra- and inter-subject variability. The impact of different strategies for voxel selection was assessed.

Results—There were no significant differences between repeated measurements in the same subject. For the 28 min protocol, PCr/ATP in the midseptal voxel across all scans was 1.91 ± 0.36 (mean ± inter-subject SD). For the 6½ min protocol, PCr/ATP in the midseptal voxel was 1.76 ± 0.40. The coefficients of reproducibility (CR) were 0.49 (28 min) and 0.67 (6½ min). Per-subject B_0-shimming improved the fitted PCr/ATP precision (for 6½ min scans), but had negligible effect on the coefficients of reproducibility (0.67 *vs* 0.66).

Conclusions—Both 7T protocols show improved reproducibility compared to a previous 3T study by Tyler *et al*. Our results will enable informed power calculations and protocol selection for future clinical research studies.

Corresponding author: Jane Ellis, jane.ellis@rdm.ox.ac.uk.

Europe PMC Funders Author Manuscripts

Keywords

^{31}P; 7T; cardiac; Magnetic resonance spectroscopy; MRS; human; reproducibility

1 Introduction

Phosphorus magnetic resonance spectroscopy (^{31}P-MRS) is a non-invasive technique used to measure the concentrations and chemical kinetics of high energy phosphorus-containing metabolites in the human heart, often collectively referred to as “cardiac energetics”. ^{31}P-MRS has provided a unique insight into our understanding of cardiac metabolism.1,2 The application of ^{31}P-MRS to cardiovascular research is of interest, since in most major heart diseases the ratio of phosphocreatine (PCr) to adenosine triphosphate (ATP) concentrations (PCr/ATP) changes, making it a useful indicator of the altered energetic state of the heart. Examples of diseases where PCr/ATP decreases include: type I3 and type II diabetes4, hypertensive heart disease5, coronary artery disease6,7 and heart failure.8,9

However, ^{31}P-MRS is yet to be translated into routine use in a clinical setting, owing mainly to its intrinsically low signal-to-noise ratio (SNR). According to theory, the quality of the raw ^{31}P-MRS signal ($SNR/\sqrt{T_A}$, where T_A is scan duration) increases approximately linearly with the scanner’s magnetic field strength B_0.10 Accordingly, ^{31}P-MRS benefits significantly from a move to ultra-high field strengths (UHF), i.e. $B_0 \geq 7T$. At 7T, an increase in SNR of 2.8× compared to 3T in the human heart has been recently demonstrated11, allowing the acquisition of spectra in 6 minutes of a comparable quality that took 31 minutes at 3T. Acquiring usable ^{31}P spectra in shorter scan times is highly desirable; a finer temporal resolution would allow study of the response cardiac energetics to stressors (e.g dobutamine infusion or exercise).

It is, however, questionable whether such a short protocol would be sufficiently robust and reproducible to detect changes in cardiac metabolism with sufficient power. Therefore, the primary motivation of this work was to assess reproducibility values of PCr/ATP ^{31}P-MRS measurements at 7T in the human heart for a ‘rapid’ 6½ minute 3D-CSI sequence. If found to be sufficiently reproducible, this would allow the study of multiple steady states within a single scanning session. We compare this to an established 28 minute protocol that was designed to record high quality spectra in a scan time that is tolerable to a majority of patients.11 This will allow informed decisions when selecting protocols in the design of future studies. We also evaluate the effect of per-subject B_0 shimming on spectral quality and reproducibility. Finally, we illustrate the impact of these technical improvements by comparing the sample sizes and approximate scan costs of studies using these optimised 3T and 7T protocols.

2 Experimental

Ten healthy volunteers (3 female, age = 29±6 years, BMI = 23±4 kg/m^2) were recruited according to local ethics regulations. Volunteers made two separate visits. On each visit they underwent two scanning sessions as shown in Figure 1. On the first visit, two 28-minute ^{31}P-

MRS spectra were acquired (one during each scan session) using the manufacturer's default shim settings. On the second visit, three 6½ minute ^{31}P-MRS spectra were acquired in both scanning sessions. During the second visit, we also tested a customised B_0 shimming algorithm against the manufacturer's default shim settings.

2.1 Data acquisition

All scans were performed on a Magnetom 7T MRI scanner (Siemens, Erlangen, Germany). To facilitate coil swapping, subjects were scanned head-first supine. A 10cm ^{1}H transmit/receive (Tx/Rx) loop coil (Rapid Biomedical, Rimpar, Germany) was used to acquire 2 chamber, 4 chamber and mid-short-axis (SA) ^{1}H fast low angle shot (FLASH) stacks of localiser cine images.

B_0 maps were acquired as a stack of 18 dual-echo GRE slices aligned with the mid-short axis view (TR 314ms, TE_1 2.4ms TE_2 4.3ms). B_0 maps were measured starting from the scanner manufacturer's default shim settings (the "tune-up" shim settings). Per-subject shim solutions were obtained by recording a B_0 map using a dual-echo gradient-recalled echo shim (GRE shim) works-in-progress package (WIP 452B "GRE SHIM", Siemens, Erlangen, Germany). This field map was altered in MATLAB by zeroing all pixels with an intensity less than that in the inferior myocardium and setting the magnitude of all remaining pixels to 1. The phases of the pixels were left unaltered. This modified B_0 map was then used as input to the scanner's standard shim calculation routines. These modifications prevent the very high intensity pixels near the surface dominating the computed shim solution, which might otherwise degrade the shim quality in the inferior cardiac segments.

The ^{1}H coil was replaced with a 16 element ^{31}P receive array coil (Rapid Biomedical, Rimpar, Germany), consisting of a single rectangular 28 × 27cm^2 transmit loop and a 4 × 4 matrix of sixteen 5.5cm-diameter circular flexible receive loops. The coil was placed in the same position above the mid-ventricular septum.12 The transmit efficiency was calibrated per-subject by using a series of inversion-recovery free induction decays (IR FIDs) to acquire signal from a central spherical phenylphosphonic acid (PPA) fiducial mounted on the coil housing and processed with custom MATLAB (The Mathworks Inc, Natick, MA, USA) code. The coil position was determined from three orthogonal single-channel ^{31}P FLASH images which localise five PPA fiducials (including the centre fiducial used to compute the B_1 calibration) using custom MATLAB code.

A 25mm thick, B_1-insensitive train to obliterate signal (BISTRO), saturation band was placed in the anterior chest wall to suppress signal from skeletal muscle as previously described.13 The voltage of the saturation pulse was set to the maximum value allowed to comply with the legal specific absorption rate (SAR) limits in each subject. Excitation was placed at +266Hz relative to PCr so as to cover metabolites from 2,3-DPG to γ-ATP. The CSI grid was positioned in the short axis view of the heart, such that the longest voxel dimension was aligned with the intraventricular septum and the in-plane voxel matrix was parallel to the chest wall. The CSI matrix was fixed at the point of acquisition, and not shifted in post-processing. Respiratory gating and ECG triggering were not used.

In the first visit, a single ^{31}P dataset was acquired in 28 minutes with a 3D UTE-CSI sequence with the following parameters: matrix size 16×16×8; nominal voxel size 15×15×25mm^3; acquisition weighting with 10 averages at k=0; repetition time 1 s and whitened singular value decomposition (WSVD) coil combination.14 The excitation was set to 400V peak amplitude (i.e. 3.2kW), giving a flip angle of approximately 30° in the interventricular septum. RF excitation was performed using a shaped pulse that has been previously described.15 It comprises a 0.5ms hard pulse, preceded by a numerically optimised 1.9ms part that improves homogeneity of excitation. It excites an approximately 2kHz bandwidth. Subjects were then removed from the magnet – this constituted 'session 1'. After a short (˜5min) break, localisation was repeated and an identical 28-minute 3D UTE-CSI sequence was run in 'session 2'. These datasets are labelled X and Y (see Fig. 1 for detailed description).

During the second visit, 3 sets of ^{31}P spectra were acquired in 6 ½ minutes each. The same 3D UTE-CSI sequence was used as described above, but with an 8×16×8 matrix; nominal voxel size 25×15×25mm^3 and 4 averages (k=0) to enable the shorter acquisition time. Using larger voxels allowed us to reduce scan duration while keeping enough samples at k=0 to preserve a compact voxel point-spread-function with minimal side-lobes. The first and third datasets used per-subject B_0 shimming; the second used the vendor's standard "tune up" shim settings. As in the first visit, subjects were then removed from the magnet – this constituted 'session 3'. After a short (˜5min) break, these steps were repeated as 'session 4'. These datasets are labelled A-F (see Fig. 1 for detailed description).

2.2 Data analysis

Four voxels in the mid-interventricular septum were identified for further analysis: the midseptal, anteroseptal, anterior and posterior voxels. These were assigned anatomically; the midseptal voxel was two voxels posterior to the chest wall. Data from these voxels were fitted using the Oxford Spectroscopy Analysis (OXSA) toolbox's implementation of AMARES.16,17 Prior knowledge specified 11 Lorentzian peaks, fixed amplitude ratios, and literature values for the scalar couplings for the multiplets. Blood contamination and partial saturation was corrected using T_1 values from the literature.11,15 All scans were included in the analysis. PCr/ATP is reported as the blood- and saturation-corrected values of PCr/γ-ATP, excluding the α-ATP peak because it has contributions from NADH and the β-ATP peak because it was outside the uniform flip-angle bandwidth of the excitation pulse at 7T. Cramér-Rao lower bounds (CRLBs) were used to express the uncertainty in metabolite concentrations.18 Reproducibility statistics of the four identified voxels and their spectral sum were calculated.

2.3 Assessment of reproducibility

Inter-session variability was assessed through the mean and difference between PCr/ATP ratios from equivalent datasets in both protocols for each subject. i.e. by comparing dataset B with E, and comparing dataset C with F, and X with Y (see Figure 1 for detailed description).

Europe PMC Funders Author Manuscripts

Inter-subject variability was assessed by the mean and standard deviation of PCr/ATP ratios within the same datasets across all subjects.

Intra-session variability, only applicable to the second visit, was assessed through the mean and difference between PCr/ATP ratios from equivalent datasets within the same session. i.e. by comparison of dataset A with C and dataset D with F.

The coefficient of repeatability (CR) was calculated from standard deviation of the *signed* differences in PCr/ATP between two scans for each subject according to:

$$CR = SD_{intrasubject} \times 1.96. \quad [1]$$

A *lower* CR reflects a better method. A two-tailed Mann-Whitney U (non-parametric) test was used to compare repeated measurements. Variances were compared using a Brown-Forsythe test.19 The coefficient of variation (CV), defined as the sample standard deviation divided by the mean, is also reported.

$$CV = SD/mean\,. \quad [2]$$

Datasets X and Y were compared to those collected from 25 patients with dilated cardiomyopathy (DCM) in a previous study using the same 28-minute protocol, coils, and scanner.20

3 Results

Typical spectra from the 28 minute and the 6½ minute CSI protocols (using the tune-up shim) from across the heart shown are shown in Figure 2. Lower signal is observed in the posterior voxels owing to the use of a surface coil. Across all subjects, the SNR of PCr was 1.2× greater in the 28 minute CSI scan compared to the 6½ minute scan (having fewer averages but larger voxels). Furthermore, in the 6 ½ minute scan we observed a larger 2,3-DPG amplitude (3.5 ± 0.62 vs 2.9 ± 1.06, $P = 0.07$), reflecting greater blood contamination from the larger voxels.

3.1 Analysis methods

Figure 3 gives reproducibility statistics for single voxels, and combinations of voxels summed prior to fitting. For both protocols, the midseptal voxel gives the most reproducible PCr/ATP values, shown by it having the lowest coefficient of reproducibility. The voxels closest to the chest wall (the anterior and anteroseptal voxels) have the highest PCr SNR, and the lowest PCr/ATP CRLB indicating a more precise quantification. However, the reproducibility in these voxels is lower than in the midseptal voxel. The posterior voxel has both a lower quantification precision (higher CRLBs) and worse reproducibility than the mid-septal voxel. Summing spectra from the midseptal voxel with anteroseptal and anterior voxels before fitting leads to a reduction in PCr/ATP CRLB but comes at a cost to the

reproducibility of the measurement. All subsequent analyses therefore calculated PCr/ATP from the midseptal voxel, not the summed voxels.

3.2 Assessment of reproducibility

28 min protocol—There were no significant differences in PCr/ATP (1.86±0.32 vs 1.96±0.39, $P = 0.79$) between repeated measurements in the same subjects (Fig. 4). The mean PCr/ATP measured across all scans was 1.91±0.36, giving a CV of 18%. The intra-subject variability was 0.21± 0.16, giving a CR of 0.49. These values are shown on Bland-Altman plots in Figure 5.

6 ½ min protocol—There were no significant differences (Fig. 4) between repeated measurements using the default “tune up” shim (1.69± 0.48 vs 1.79±0.30, $P = 0.57$), or the customised shimming (1.64±0.26 vs 1.65±0.17, $P = 0.97$). Mean PCr/ATP across all scans was 1.76±0.40 for the tune-up and 1.70±0.28 for the custom shim, giving CVs of 23% and 17% respectively. The intra-subject variabilities were 0.31±0.14 (tune-up) and 0.29±0.18 (per-subject shim). There was an improvement in the Cramér-Rao lower bounds (CRLBs) between tune-up and the customised shimming algorithms (13.8% vs 10.3%, $P = 0.02$), showing that the higher quality spectra obtained using the customised shimming enabled more precise metabolite quantification. The reproducibilty of PCr/ATP measurements using both shimming techniques was equivalent (CR: 0.66 tune-up shim vs 0.67 custom shim).

The inter-examination CR was 0.75 and the intra-examination CR was 0.67 (both reported for the custom shim comparison). There was no significant difference (P = 0.85) between these inter- and intra-examination changes.

3.3 Power calculations and sample size

Sample sizes providing sufficient power to reveal differences in PCr/ATP for paired (e.g. response to stressors or patient longitudinal cohort) studies are given in Table 1. These were calculated using the reproducibility values obtained in this work and literature results at 3T.

Discussion

We have tested the reproducibility of a ‘rapid’ 6½ minute 3D-CSI protocol, which is short enough to allow the assessment of changes in PCr/ATP during rest, stress and recovery in a single scanning session. We also tested reproducibility for an established 28 minute 3D-CSI protocol that was designed to give the best spectra in an examination tolerable to cardiac patients. We believe this study reports the first reproducibility data for human cardiac ^{31}P-MRS at 7T. These are crucial for planning future clinical studies.

Reproducibility of human cardiac ^{31}P-MRS has previously been reported at lower field strengths (summarised in Table 2). These studies encompass a range of field strengths, localisation methods and voxel volumes, which makes it hard to draw quantitative conclusions when comparing these studies. However, it is clear that our 28-minute protocol has the tightest reproducibility of all the 3D-resolved protocols, and that it achieves this with a small 5.6mL nominal voxel volume.

Europe PMC Funders Author Manuscripts

4.1 Analysis methods

The midseptal voxel is the most reproducible of all the single voxels and gives PCr/ATP consistent with literature values. The anterior and anteroseptal voxels have higher PCr/ATP values and we hypothesise that this is due to contamination of the "cardiac" spectra by small amounts of skeletal muscle signal that may not be fully suppressed in some scans by the saturation bands (PCr/ATP is approximately 4-5 in skeletal muscle vs 2 in myocardium).21 Quantification of PCr/ATP values in the posterior voxels is the least reproducible, and for the 6½ minute protocol has an extremely large CRLB (111%). This is likely due to a combination of the effects of distance from the surface coil, and of motion in the posterior voxel.

In 2008, Tyler et al. performed a cardiac ^{31}P-MRS reproducibility study at 3T.15 They used the spectral sum from 3 voxels in their analysis because they observed lower PCr/ATP CRLB than for a single midseptal voxel. They attributed this increase in precision to the gain in SNR from combining signals. In this study, at 7T, we also saw increased SNR on summing voxels, but no corresponding improvement in reproducibility.

Our analysis method introduces minimal user bias: no spectra were excluded from the study based on appearance and as the rest of the fitting in OXSA is automated, measured PCr/ATP is only dependent on which voxel was selected for analysis. However, as measured PCr/ATP varied at differing anatomical locations across the heart, the selection of the voxel is an important step. It was therefore important to define anatomically which voxel will be selected for analysis at the start of the study, and ensure that it was used in all subjects.

4.2 Assessment of reproducibility

28 min protocol—Measurements reported here are consistent with literature values.11 While repeated measurements of PCr/ATP measured here showed no significant differences to each other, PCr/ATP measured in this work did show significant difference to values measured in 25 DCM patients using an identical 28-minute CSI protocol and the same hardware (1.91±0.36 vs 1.54±0.39, $P = 0.01$).9

Despite SNR limitations, 3T ^{31}P-MRS has allowed observation of cardiac energetic changes in many disease groups. In recent years, 3T ^{31}P-MRS has been the most widely used field strength for cardiac phosphorus scans. The 28 minute CSI protocol tested here has equal voxel sizes (5.6mL nominal) and a similar scan duration (28 vs 31 minutes) to the 3T protocol tested by Tyler et al.15 In that study, they found intra-subject percentage differences of 20% (absolute difference = 0.43, mean PCr/ATP = 2.10) and a coefficient of reproducibility of 1.1. In this study, we report lower intra-subject percentage differences of 11% (absolute difference = 0.21, mean PCr/ATP = 1.91) and a coefficient of reproducibility of 0.49 (i.e. more reproducible).

6½ minute protocol—The 'rapid' 6½ minute CSI acquired datasets in less than ¼ of the time to both those in Tyler et al.'s 3T study and the 28 minute scan tested here. As expected, a decrease in scan time led to increased variability in PCr/ATP measurements and therefore a larger CR (0.49 28 minute CSI vs 0.67 6½ minute CSI). Despite this, the reproducibility of

PCr/ATP measurements acquired in 6½ minutes at 7T is still greater than those acquired in 31 minutes at 3T, meaning that a move to ultra-high field means that data can be acquired quicker without cost to reproducibility. This is important as short ^{31}P acquisition times allow the detailed study of the response of cardiac energetics to stressors, such as exercise or dobutamine infusion. For example, the British Society of Echocardiography's (BSE) recommendations for a dobutamine protocol for assessment of myocardial ischemia are to use four steps each with 3 minutes of dobutamine infusion.22 7T ^{31}P stress using our 6½ min ^{31}P scan would therefore be possible during dobutamine stress complying with these guidelines, whereas a 31 minute scan at 3T would not be. We note in passing that Dass et al proposed a shorter 8 minute ^{31}P scan at 3T, and demonstrated that it gave similar mean PCr/ATP values to the Tyler et al 31 minute 3T protocol (Dass et al 8min: 1.83 ± 0.32; Tyler et al 31min: 1.78 ± 0.27).23 However, the reproducibility of that 8min 3T protocol has not been reported, so we cannot compare its reproducibility against our 7T results.

Bakermans et al tested the reproducibility of a 7 minute sequence in the human heart at 3T, performing spatial localisation with 3D ISIS or a combination of 1D ISIS and 1D CSI either perpendicular or parallel to the surface coil.24 3D ISIS was found to be the most reproducible of these methods, giving a PCr/ATP of 1.57 ± 0.17 (mean ± SD) and a CR of 0.64. The reproducibility of this 7 min 3T scan is the same as that achieved in a 6½ min 7T scan in this work (CR 0.64 3T vs CR 0.67 this work) but used substantially larger voxel sizes: 3D ISIS voxel size 512 mL. We observed no significant differences between the variances of the intra- and inter- examination differences ($P = 0.85$), suggesting that the variability is measured PCr/ATP is dominated by error from within the ^{31}P measurement itself, rather than experimental set-up (e.g. coil positioning, localisation). This is in contrast to the finding of Lamb et al. in their reproducibility study at 1.5T.25 There, they used a 10cm ^{31}P loop for both transmission and reception. In this study we used a bigger coil with a rectangular 26x28cm^2 transmit loop and a flexible 4x4 array of 4cm diameter receive elements. By using a larger coil, we have mitigated some of the challenges associated with placing small loop coils and so our results were less affected.

Per-subject B_0 shimming

Per-subject B_0 shimming improved the precision of the fit of PCr/ATP, as indicated by the lower CRLBs (mean 13.8% tune-up vs 10.3% per-subject, $P = 0.02$). However, on our system, the reductions in the linewidths of PCr and γATP were not significant and there were no improvements in reproducibility (CR 0.67 tune-up, CR 0.66 per-subject). Presumably uncertainties in optimising the shim solution counter-balanced the improved spectral SNR in terms of reproducibility. Additionally, as no cardiac triggering or respiratory gating was used, the calculated shims only apply exactly at one phase in the cardiac and respiratory cycle, which might explain the lack of improvement. It is not therefore immediately obvious whether per-subject B_0 shimming for cardiac ^{31}P-MRS at 7T is worth the additional examination time that is required: two iterations of the shimming algorithm adds two approximately 20 second breath holds to the protocol.

If a single 6½ minute ^{31}P measurement is being included in an examination, then it is likely that per subject B_0 shimming is not worth the extra time. In this case, if time permits, better

data would probably be obtained by using the time that would have been spent shimming to acquire more averages in the CSI protocol. However, if stress-response energetics are being monitored (and therefore a CSI protocol is being repeated multiple times while the subject remains in the magnet) then all of the datasets would benefit from performing the per-subject shimming algorithm at the start of the scan.

Our difficulties with image-based shimming at 7T may stem from the inhomogeneous fields produced by the 10cm surface coil we used. More sophisticated coil designs with better coverage, and dual-tuned designs may tip the balance in favour of per-subject B_0 shimming.[26]

4.3 Power and sample size

The reproducibility values presented here enable power and sample size calculations to be performed — an important step in the design of clinical studies. The lower intra-subject variability at 7T compared to 3T translates to smaller sample sizes required for sufficient statistical power to detect a given effect. For example, in order to detect a change of 0.2 in PCr/ATP with 80% power in a paired study, power calculations from 3 T data from the Tyler et al.[15] study predict a required sample size of 64, whereas data from this work predicts only 15 subjects would be required for an equivalent length protocol.

There is some debate in the literature about the possible magnitude of changes in PCr/ATP. Bakermans *et al.* suggested by simulations that the changes in PCr/ATP in healthy subjects upon exercise are on the order of 10%, which they concluded would be impossible to detect by ^{31}P-MRS at 3T. Exercise-induced changes in patients with disease are known to be more substantial, perhaps because their hearts have only a limited energetic reserve because of having to compensate for disease at rest. For example, Betim Paes Leme *et al.* observed a 57% decrease in PCr/ATP in patients with Chagas heart diseaseon exercise.[27] Such changes would be clearly detectable at 7T. It is an open question whether 7T ^{31}P-MRS can be made sensitive enough to detect changes in high-energy phosphate energetics in healthy subjects during exercise.

When designing studies, cost is an important factor. Scan fees at 7T are more expensive compared to an equivalent scan at 3T: for example, in our centre, the cost of a 7T scan is ~70% more than a 3T scan. Despite this, the smaller cohort size required at 7T for sufficient statistical power would overall lead to a saving in scan fees, e.g. a 100 x (1 –1.70 x 15/64) = 60% saving, using values for the 28 minute protocol. Additionally, there would be further savings from the use of fewer consumables and less time spent on patient recruitment. This may also enable a shorter study, delivering more timely information to clinical decision makers. However, at present, our site has more restrictive exclusion criteria for scans at 7T than at 3T. Nevertheless, in a recent 7T ^{31}P-MRS study in patients with dilated cardiomyopathy[20] approximately 60% of patients who completed the laboratory screening form were found to have no safety contraindication to MR imaging at 7T and were able to participate fully. In our experience, the challenges of patient safety are not insurmountable at 7T. In particular, we would have excluded many fewer subjects if we had had access to more complete medical device testing data that included 7T. We expect this will become available in the coming years.

4.4 Limitations

We analysed our results in terms of the PCr/ATP ratio, which is the most commonly measured parameter in cardiac ^{31}P-MRS. However, the PCr/ATP ratio is typically used under the assumption that [PCr] changes while [ATP] remains constant. This assumption is reasonable in the healthy heart, but at high workloads [ATP] decreases and so using the PCr/ATP ratio may obscure changes in ATP concentration.1 In future work, methods such as absolute quantitation (i.e the concentration of metabolites are calibrated to recognised units e.g. mol/L) would overcome this limitation.28

The aim of this work was to assess reproducibility of cardiac PCr/ATP at 7T in the healthy population. At 7T the transmit field strength B_1^+ – and therefore the flip angle achieved in the myocardium – depends strongly on coil loading. Reproducibility may therefore be different in patient groups whose body shape is different to the healthy volunteers in this study (e.g obese type 2 diabetes mellitus patients).

Conclusion

We report reproducibility values for human cardiac ^{31}P-MRS at 7T. These provide the necessary information to design future clinical studies using 7T ^{31}P-MRS as an endpoint. We evaluated two protocols, one 28 min CSI protocol designed to acquire the best quality spectra in a clinically feasible scan time, and one 6½ min CSI protocol that gives spectra of a quality previously reported at 3T,15 but in a time short enough for use in stress-response studies. Per-subject B_0 shimming improved spectral quality, but had a negligible impact on measurement reproducibility.

Acknowledgements

We thank Professor D. J. Tyler for providing the raw data from his 3 T reproducibility study (ref 15). Funded by a Sir Henry Dale Fellowship from the Wellcome Trust and the Royal Society to CTR [098436/Z/12/B]. JE receives a DPhil studentship from the Medical Research Council (UK). The support of the Slovak Grant Agency VEGA (grant #2/0001/17) and APVV (grant #15-0029) is also acknowledged.

List of abbreviations

ATP	adenosine triphosphate
CR	Coefficient of reproducibility
CSI	Chemical shift imaging
CV	coefficient of variation
DCM	dilated cardiomyopathy
FLASH	fast low angle shot
OXSA	Oxford Spectroscopy Analysis
PCr	Phosphocreatine
PPA	phenylphosphonic acid

UHF	ultra high field
UTE-CSI	Ultra short echo chemical shift imaging
WSVD	whitened singular value decomposition

References

1. Neubauer S. The failing heart--an engine out of fuel. N Engl J Med. 2007; 356(11):1140–1151. DOI: 10.1056/NEJMra063052 [PubMed: 17360992]
2. Hudsmith LE, Neubauer S. Detection of myocardial disorders by magnetic resonance spectroscopy. Nat Clin Pract Cardiovasc Med. 2008; 5:S49–S56. DOI: 10.1038/ncpcardio1158 [PubMed: 18641607]
3. Metzler B, Schocke MFH, Steinboeck P, et al. Decreased high-energy phosphate ratios in the myocardium of men with diabetes mellitus type I. J Cardiovasc Magn Reson. 2002; 4(4):493–502.Toegang verkry Junie 13, 2017 [PubMed: 12549236]
4. Scheuermann-Freestone M, Madsen PL, Manners D, et al. Abnormal Cardiac and Skeletal Muscle Energy Metabolism in Patients With Type 2 Diabetes. Circulation. 2003; 107(24):3040–3046. DOI: 10.1161/01.CIR.0000072789.89096.10 [PubMed: 12810608]
5. Heyne J-P, Rzanny R, Hansch A, Leder U, Reichenbach JR, Kaiser WA. 31P-MR spectroscopic imaging in hypertensive heart disease. Eur Radiol. 2006; 16(8):1796–1802. DOI: 10.1007/s00330-006-0170-0 [PubMed: 16514468]
6. Neubauer S, Krahe T, Schindler R, et al. 31P magnetic resonance spectroscopy in dilated cardiomyopathy and coronary artery disease. Altered cardiac high-energy phosphate metabolism in heart failure. Circulation. 1992; 86(6)Toegang verkry Junie 13, 2017
7. Yabe T, Mitsunami K, Inubushi T, Kinoshita M. Quantitative Measurements of Cardiac Phosphorus Metabolites in Coronary Artery Disease by 31P Magnetic Resonance Spectroscopy. Circulation. 1995; 92(1)Toegang verkry Junie 13, 2017
8. Weiss RG, Gerstenblith G, Bottomley PA. ATP flux through creatine kinase in the normal, stressed, and failing human heart. Proc Natl Acad Sci. 2005; 102(3):808–813. DOI: 10.1073/pnas.0408962102 [PubMed: 15647364]
9. Neubauer S, Horn M, Pabst T, et al. Contributions of 31P-magnetic resonance spectroscopy to the understanding of dilated heart muscle disease. Eur Heart J. 1995; 16(Suppl O):115–118.Toegang verkry Junie 12, 2017 [PubMed: 8682076]
10. Redpath TW. Signal-to-noise ratio in MRI. Br J Radiol. 1998; 71(847):704–707. DOI: 10.1259/bjr.71.847.9771379 [PubMed: 9771379]
11. Rodgers CT, Clarke WT, Snyder C, Vaughan JT, Neubauer S, Robson MD. Human cardiac ^{31}P magnetic resonance spectroscopy at 7 tesla. Magn Reson Med. 2014; 72(2):304–315. DOI: 10.1002/mrm.24922 [PubMed: 24006267]
12. Rodgers, CT; Clarke, WT; Berthel, D; Neubauer, SRM. A 16-element receive array for human cardiac 31P MR spectroscopy at 7T. Proceedings of the 22th Annual Meeting of ISMRM; Milan, Italy. 2014. 22nd Annual Meeting of ISMRM
13. Luo Y, de Graaf RA, DelaBarre L, Tannús A, Garwood M. BISTRO: An outer-volume suppression method that tolerates RF field inhomogeneity. Magn Reson Med. 2001; 45(6):1095–1102. DOI: 10.1002/mrm.1144 [PubMed: 11378888]
14. Rodgers CT, Robson MD. Receive array magnetic resonance spectroscopy: Whitened singular value decomposition (WSVD) gives optimal Bayesian solution. Magn Reson Med. 2010; 63(4):881–891. DOI: 10.1002/mrm.22230 [PubMed: 20373389]
15. Tyler DJ, Emmanuel Y, Cochlin LE, et al. Reproducibility of ^{31}P cardiac magnetic resonance spectroscopy at 3 T. NMR Biomed. 2009; 22(4):405–413. DOI: 10.1002/nbm.1350 [PubMed: 19023865]
16. Purvis LAB, Clarke WT, Biasiolli L, Valkovič L, Robson MD, Rodgers CT. OXSA: An open-source magnetic resonance spectroscopy analysis toolbox in MATLAB. Motta A, red. PLoS One. 2017; 12(9):e0185356.doi: 10.1371/journal.pone.0185356 [PubMed: 28938003]

17. Vanhamme L, van den Boogaart A, Van Huffel S. Improved Method for Accurate and Efficient Quantification of MRS Data with Use of Prior Knowledge. J Magn Reson. 1997; 129(1):35–43. DOI: 10.1006/JMRE.1997.1244 [PubMed: 9405214]
18. Cavassila S, Deval S, Huegen C, van Ormondt D, Graveron-Demilly D. Cramér-Rao bounds: an evaluation tool for quantitation. NMR Biomed. 2001; 14(4):278–283. DOI: 10.1002/nbm.701 [PubMed: 11410946]
19. Brown MB, Forsythe AB. Robust Tests for the Equality of Variances. J Am Stat Assoc. 1974; 69(346):364–367. DOI: 10.1080/01621459.1974.10482955
20. Stoll VM, Clarke WT, Levelt E, et al. Dilated Cardiomyopathy: Phosphorus 31 MR Spectroscopy at 7 T. Radiology. 2016; 281(2):409–417. DOI: 10.1148/radiol.2016152629 [PubMed: 27326664]
21. Kemp GJ, Meyerspeer M, Moser E. Absolute quantification of phosphorus metabolite concentrations in human muscle in vivo by 31P MRS: a quantitative review. NMR Biomed. 2007; 20(6):555–565. DOI: 10.1002/nbm.1192 [PubMed: 17628042]
22. BSE procedure guidelines for the clinical application of stress echocardiography, recommendations for performance and interpretation of stress echocardiography. Heart. 2004; 90:23–30. DOI: 10.1136/hrt.2004.047985
23. Dass S, Cochlin LE, Holloway CJ, et al. Development and validation of a short 31P cardiac magnetic resonance spectroscopy protocol. J Cardiovasc Magn Reson. 2010; 12(Suppl 1):P123.doi: 10.1186/1532-429X-12-S1-P123
24. Bakermans AJ, Bazil JN, Nederveen AJ, et al. Human Cardiac 31P-MR Spectroscopy at 3 Tesla Cannot Detect Failing Myocardial Energy Homeostasis during Exercise. Front Physiol. 2017; 8:939.doi: 10.3389/fphys.2017.00939 [PubMed: 29230178]
25. Lamb HJ, Doornbos J, den Hollander JA, et al. Reproducibility of Human Cardiac31P-NMR Spectroscopy. NMR Biomed. 1996; 9(5):217–227. DOI: 10.1002/(SICI)1099-1492(199608)9:5<217::AID-NBM419>3.0.CO;2-G [PubMed: 9068003]
26. Valkovič L, Dragonu I, Almujayyaz S, et al. Using a whole-body 31P birdcage transmit coil and 16-element receive array for human cardiac metabolic imaging at 7T. Lundberg P, red. PLoS One. 2017; 12(10):e0187153.doi: 10.1371/journal.pone.0187153 [PubMed: 29073228]
27. Betim Paes Leme AM, Salemi VMC, Weiss RG, et al. Exercise-Induced Decrease in Myocardial High-Energy Phosphate Metabolites in Patients With Chagas Heart Disease. J Card Fail. 2013; 19(7):454–460. DOI: 10.1016/J.CARDFAIL.2013.05.008 [PubMed: 23834921]
28. Beer M, Seyfarth T, Sandstede J, et al. Absolute concentrations of high-energy phosphate metabolites in normal, hypertrophied, and failing human myocardium measured noninvasively with 31P-SLOOP magnetic resonance spectroscopy. J Am Coll Cardiol. 2002; 40(7):1267–1274. DOI: 10.1016/S0735-1097(02)02160-5 [PubMed: 12383574]
29. Abozguia K, Elliott P, McKenna W, et al. Metabolic Modulator Perhexiline Corrects Energy Deficiency and Improves Exercise Capacity in Symptomatic Hypertrophic Cardiomyopathy. Circulation. 2010; 122(16):1562–1569. DOI: 10.1161/CIRCULATIONAHA.109.934059 [PubMed: 20921440]
30. Schaefer S, Schwartz GG, Steinman SK, Meyerhoff DJ, Massie BM, Weiner MW. Metabolic response of the human heart to inotropic stimulation:In vivo phosphorus-31 studies of normal and cardiomyopathic myocardium. Magn Reson Med. 1992; 25(2):260–272. DOI: 10.1002/mrm.1910250205 [PubMed: 1614310]

Europe PMC Funders Author Manuscripts

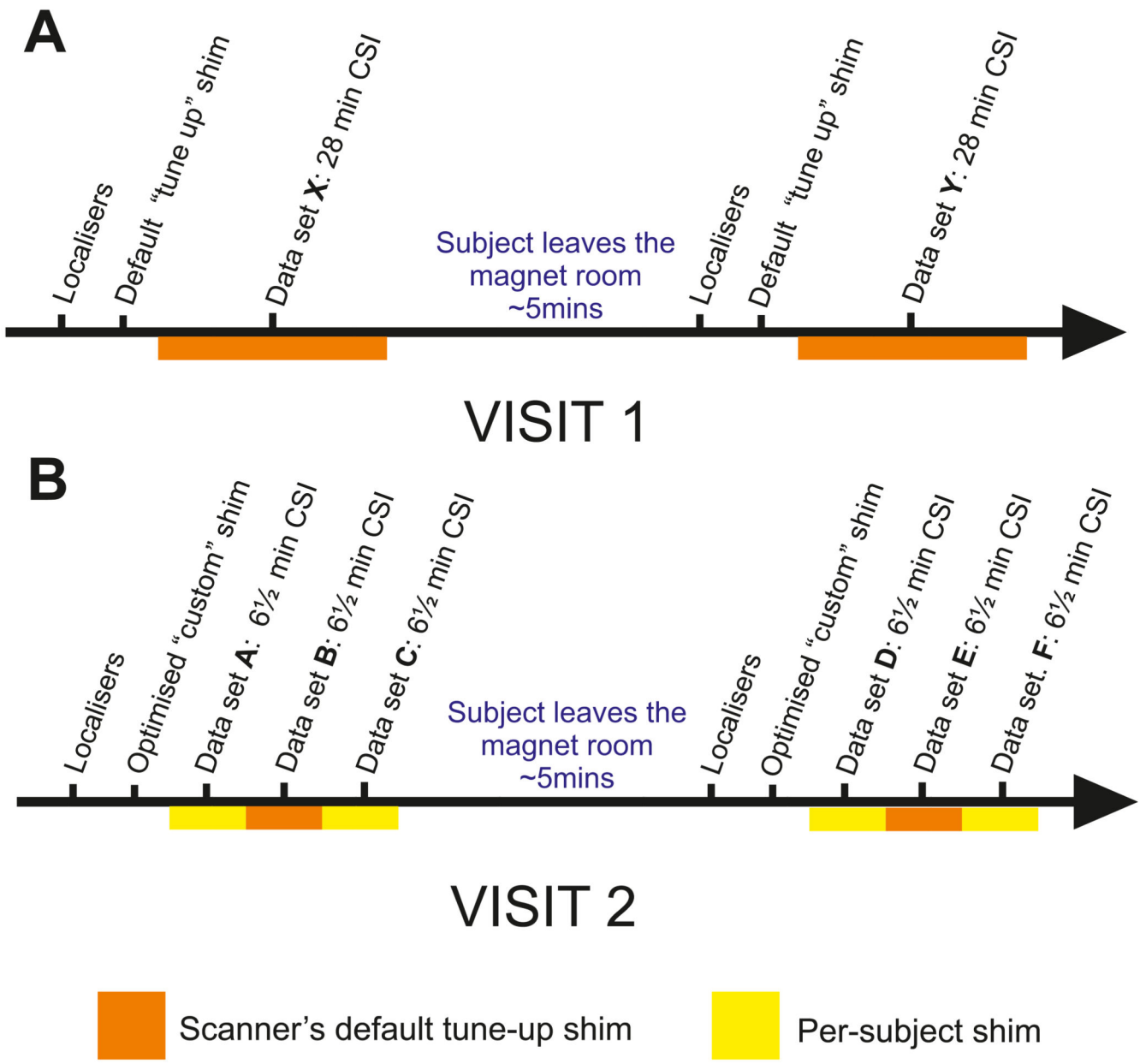


Figure 1.

Study protocol. A 'session' was defined as a series of consecutive sequences, acquired without the patient leaving the magnet. 'Dataset' refers to a single CSI sequence acquired within a session. Panel **A** shows the visit protocol for acquisition of 28 min CSI datasets, **B** the 6½ min datasets.

Europe PMC Funders Author Manuscripts

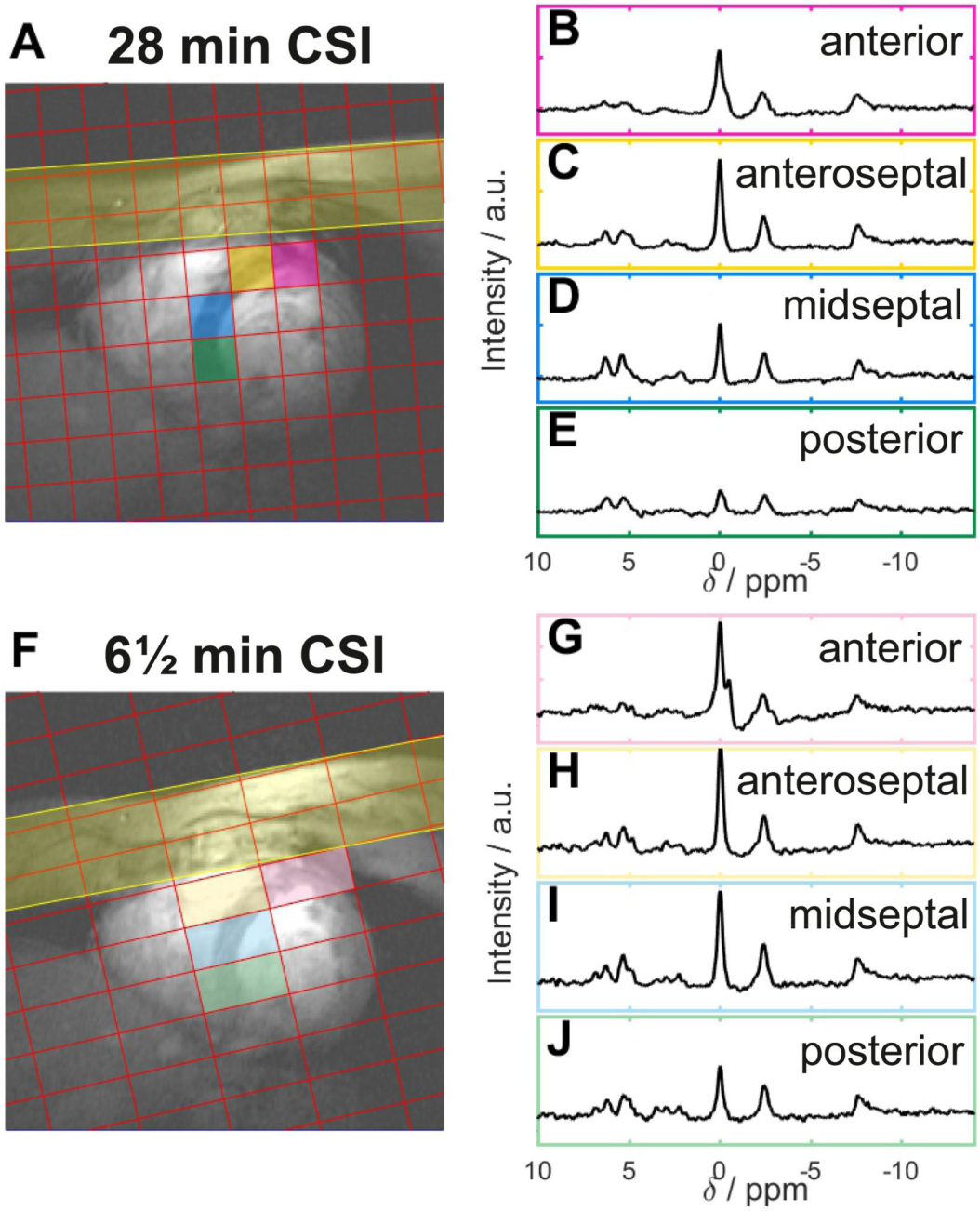
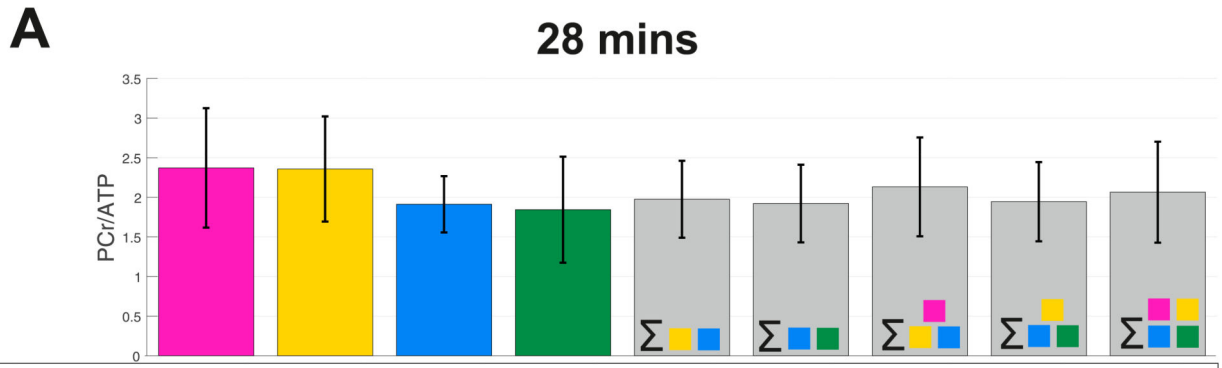
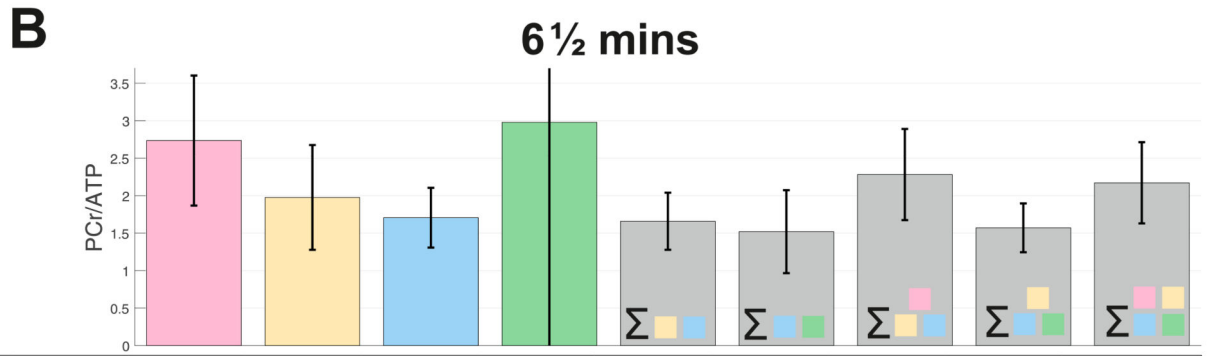


Figure 2.

A, F Positions of the CSI matrix showing the rotation of the CSI grid in the short axis view of the heart for both protocols. **B-E,G-J** Spectra from the corresponding coloured voxel marked on the localiser images



	Anterior voxel	Anteroseptal voxel	Midseptal voxel	Posterior voxel	Anteroseptal + Midseptal	Midseptal + Posterior	Anterior + Anteroseptal + Midseptal	Anteroseptal + Midseptal + Posterior	All 4 voxels
PCr/ATP	2.37	2.35	1.91	1.84	1.98	1.92	2.13	1.95	2.06
PCr SNR	39	41	28	21	50	34	74	55	77
CRLB	9.1%	8.1%	9.9%	15%	8.8%	13%	8.6%	10%	9.8%
Coefficient of Variation	32%	28%	18%	37%	25%	26%	23%	26%	31%
Coefficient of Reproducibility	2.0	1.2	0.49	2.3	1.23	1.42	1.19	1.47	1.19



	Anterior voxel	Anteroseptal voxel	Midseptal voxel	Posterior voxel	Anteroseptal + Midseptal	Midseptal + Posterior	Anterior + Anteroseptal + Midseptal	Anteroseptal + Midseptal + Posterior	All 4 voxels
PCr/ATP	2.74	1.98	1.71	2.98	1.66	1.52	2.28	1.57	2.17
PCr SNR	33	32	23	8	50	23	72	55	79
CRLB	12%	9.3%	13%	111%	9.3%	20%	10%	10%	11%
Coefficient of Variation	87%	70%	38%	539%	38%	36%	64%	33%	54%
Coefficient of Reproducibility	1.9	1.5	0.67	19.02	1.01	1.5	1.5	0.86	1.4

Figure 3.

Variation in PCr/ATP value measured at 4 voxel locations and combinations of spectral sums of these voxels for **A** 28 minute CSI and **B** 6½ minute CSI. The tables below provide the PCr SNR, the Cramér Rao Lower Bound on PCr (CRLB), the coefficient of variation (CV = inter-subject SD / mean) and the coefficient of reproducibility (CR = 1.96 x inter-scan SD) of the measured PCr/ATP. Please note, the values here refer to inter-examination repeatability.

Europe PMC Funders Author Manuscripts

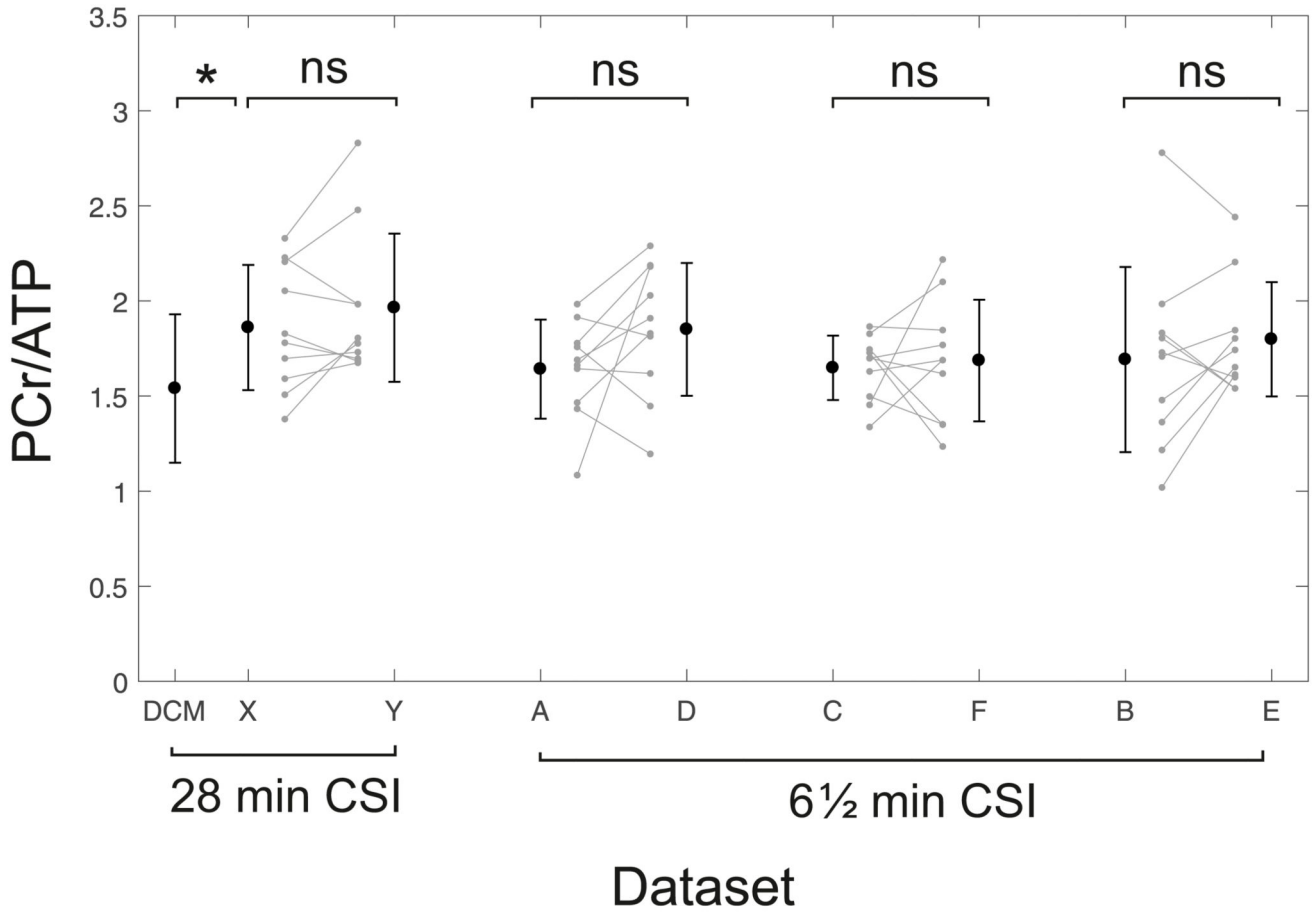


Figure 4.

Inter-subject variability between scans for the 28 minute CSI protocol (left) and the 6½ minute CSI protocol (right). Error bars show ±1x SD. The left-most bar shows DCM patient data acquired by Stoll et al20 using the same scanning protocol and hardware for comparison.

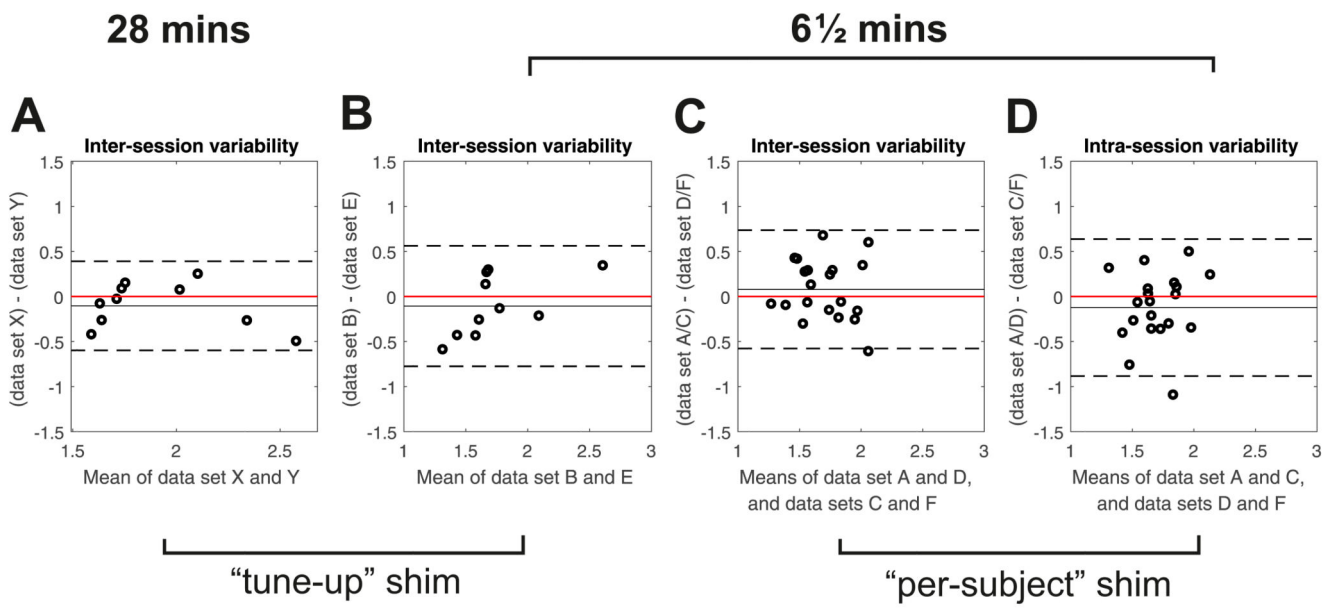


Figure 5.

Bland-Altman plots of inter-session variability in PCr/ATP for **A** 28 min CSI and **B,C** 6½ min CSI. **D** Plot of intra-session variability for 6½ min CSI with per-subject shimming. Solid black lines show the bias from zero (red line), dashed lines mark ±1.96SD.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Table 1

Power calculations showing the sample size needed in a paired study to determine statsitcal significance (α = 0.05) for both 28 min and 6.5 min CSI sequences at 7T, and 31 minute CSI sequence at 3T, for a change of 0.2 in the measured PCr/ATP.

ΔPCr/ATP = 0.2	**7T**		**3T***
	28 minutes σ=0.25, α=0.05	**6½ minutes σ=0.34, α=0.05**	**31 minutes σ=0.56, α=0.05**
Power = 95%	23	40	104
Power = 90%	19	33	85
Power = 80%	15	25	64

* 3T reproducibility values for this power calculation were taken from the study by Tyler *et al.*, (ref 15). We made this analysis using the original raw data from that study provided by Prof Tyler. If comparing this table to ref 15, please note that the values reported in Table 1 of ref. 15 were computed using the standard deviation of the absolute difference in PCR/ATP for scan 1 and scan 2, whereas here we have used the standard deviation of the signed difference.

Europe PMC Funders Author Manuscripts

Table 2

Reproducibility values of human cardiac ^{31}P-MRS from the literature

Reference	Field strength (T)	Cohort size	Acquisition time (min)	Localisation method	Nominal voxel size (cm^3)	Absolute reproducibility (%)[a]
This study	7	10	28	3D-CSI	1.5×1.5×2.5 = 5.6mL	26
			6½	3D-CSI	2.5×1.5×2.5 = 9.4mL	39
Bakermans et al, 201724	3	8	7	Single voxel 3D ISIS	$(8.0)^3$ = 512mL	40
				1D ISIS with 1D CSI	-	43
				1D ISIS with 1D CSI	-	44
Tyler et al, 200915	3	20	30	3D CSI	1.5×1.5×2.5 = 5.6mL	53
Shivu et al, 201029	3	1[b]	23	3D ISIS	4.4×5.5×3.7 = 90mL	31
Schaefer et al, 199230	1.5	7	7	1D CSI	-	22
Lamb et al, 199625	1.5	16	10	3D ISIS	6.0×7.0×7.0 = 294mL	38 [c]
			15	1D CSI	-	-
			30	2D ISIS + 1D CSI	1.0cm thick slices	46 [c]

[a] Absolute reproducibility is calculated as (CR ÷ Mean PCr/ATP) × 100%

[b] A single subject was scanned 8 times to assess reproducibility.

[c] Absolute reproducibility was computed by extracting raw values from Figure 5 in ref. 25 using WebPlotDigitizer (https://apps.automeris.io/wpd/)