**Editorial comments:**

* ***NOTE: Please download this version of the Microsoft word document (File name: 54977\_R2\_060916) for any subsequent changes. Please keep in mind that some editorial changes have been made prior to peer review.***

* **Please keep the editorial comments from your previous revisions in mind as you revise your manuscript to address peer review comments. For instance, if formatting or other changes were made, commercial language was removed, etc., please maintain these overall manuscript changes.**

**•Formatting:**

**-Please do not use duplicate sentences in the short and long abstracts.**

No sentences are duplicated in the short and long abstracts.

**-1.2 – Please make 31 superscript.**

We have corrected this line to read: “Plug the 31P coil into the in-table coil connector at the end of the exam table closest to the bore.”

**-4.10 – Which step? Is this step 4.6?**

The instruction refers to step 4.5; the text now states this more clearly: “Post-exercise T1 image: repeat the pre-exercise axial T1 image (step 4.5) using the same acquisition parameters.”

**-Please include spaces between all bullet points in the protocol. See 4.11.1 and 4.11.2.**

We have included spaces between all steps and sub-steps.

**-Please use size 12 font. Parts of the results are in a different font and size from the rest of the document.**

The font and size have been corrected to Calibri, size 12 for the entire document.

**-All latin phrases (ie in vivo, in vitro) should be in italics.**

All latin phrases have been changed to italics

**•Grammar:**

**-Please make sure all sentences end with a period (see 1.1, 4.11.2, 4.11.3).**

All sentences end with a period.

**-3.1 – “Explain the subject”**

This line has been corrected and now reads: “Explain to the subject that the exercise protocol consists of three phases: an initial baseline phase, a short intense exercise phase, and a recovery phase.”

**-3.4 – “PCR”**

This has been corrected and now reads : “Observe the **PCr** peak height in the acquisition viewer window, and also view it upon completion of the exercise sequence.”

**-4.10 – Please clarify – the phrase at the beginning of the sentence seems out of place.**

This instruction has been clarified and now reads: “Post-exercise T1 image: repeat the pre-exercise axial T1 image (step 4.5) using the same acquisition parameters.”

**-Line 547 – “?MOXPHOS”**

This line has been corrected “OXPHOS”.We also have replaced ‘mOXPHOS’ with ‘OXPHOS’ throughout, to be more consistent with widely used and accepted terminology.

**•Additional detail is required:**

**-2.6 – Where is the baby oil placed?**

This instruction has been clarified: “**Secure the baby oil to the medial aspect of the thigh** with the same straps used to secure the coil to the leg. This facilitates scan localization.

”

**-4.1, 4.2 – How does one acquire this? Are these steps performed by software? If so, how?**

**4.2.1 – How is the orientation adjusted?**

**-4.4.1 – How does one place the shim box?**

**4.4.2 – How are images used for this? What should one see to identify the “sensitive region?” Then how is the shim box adjusted? Is it moved or are the coordinates adjusted?**

**-4.4.3.1 – How is a spectrum obtained?**

**-4.6.1 – What example settings are used here?**

**Comments on steps 4.1 to 4.6.1:** These are very basic steps in MR scanning, and can be easily performed without further clarification by anyone with nominal technical training allowing use an MR scanner. We expect those without prior MR experience will perform these experiments under the supervision of a trained MR technician.

**-4.7.1 – How long does the subject kick here? Or are they actually beginning to kick in the next step? This is not clear – actions should be mentioned in the order in which they occur.**

This is a rest step, and the subject does not kick during this time. The instructions have been clarified in the manuscript: “**Rest:** Apply the shim settings from the previous scan and set the sequence to acquire 20 measurements. Instruct the subject to begin kicking after a countdown. Subject remains at rest for 2 measurements.”

**•Results: Please define the error bars (SD, SEM, etc.) in Figure 8.**

The caption for Figure 8 is now: “Comparison of PCr recover times demonstrates sequentially poorer mitochondrial oxidative capacity in control subjects, non-diabetic and diabetic subjects. Error bars represent standard deviation. ”

* **If your figures and tables are original and not published previously, please ignore this comment. For figures and tables that have been published before, please include phrases such as “Re-print with permission from (reference#)” or “Modified from..” etc. And please send a copy of the re-print permission for JoVE’s record keeping purposes.**

* **JoVE reference format requires that the DOIs are included, when available, for all references listed in the article. This is helpful for readers to locate the included references and obtain more information. Please note that often DOIs are not listed with PubMed abstracts and as such, may not be properly included when Safiyaciting directly from PubMed. In these cases, please manually include DOIs in reference information.**

* **IMPORTANT: Please copy-edit the entire manuscript for any grammatical errors you may find. The text should be in American-English only. This editing should be performed by a native English speaker (or professional copyediting services) and is essential for clarity of the protocol and the manuscript. Please thoroughly review the language and grammar prior to resubmission. Your JoVE editor will not copy-edit your manuscript and any errors in your submitted revision may be present in the published version.**

* **NOTE: Please include a line-by-line response letter to the editorial and reviewer comments along with the resubmission.**

**Reviewers' comments:**

**Reviewer #1:**

***Manuscript Summary:* This manuscript addresses the need for non-invasive techniques to assess mitochondrial function and metabolism. 31P MRS has a long pedigree in the context of exercise physiology and has improved our understanding of healthy muscle energetics. Alternatives for painful biopsy are needed, particularly for patients with genetic or acquired myopathies who often require longitudinal follow-up to assess the effectiveness of therapies on muscle metabolism. While not new, the protocol suggested by the authors offers a relatively straightforward approach for introducing 31P MRS into the clinic. The authors provide reproducibility data from healthy volunteers; the clinical feasibility of the technique is demonstrated in small groups of patients with genetic (Friedrich's ataxia) or acquired (metabolic syndrome) pathologies.**

***Major Concerns:* None.**

***Minor Concerns:***

**-1. Materials and Preparation**

**1.1) (pg 5, line 216): "prior to the experiment" rather than "prior to experiment"**

This line (line 256) has been corrected: “Ensure that all necessary materials are available prior to **the** experiment. (Figure 2)”

**1.4) (pg 6, line 224): "farthest" rather than "furthest"**

This line (line 264) has been corrected: **“Place a head pillow at the other end of the MR exam table, farthest from the bore, for subject comfort.”**

**2. Subject Positioning**

**2.1) (pg 6, line 228): "lie supine" rather than "lay" (this should be checked throughout the manuscript, eg line 258)**

We have corrected this throughout the manuscript.

**-2.9) (pg 6, line 251): "centering light" might not be the conventional term here; "laser light guide" might be preferable?**

We have corrected this. The line (line 291) now reads: “Use the **laser light guide** to delineate the center of the coil, and move the table using this centering landmark to magnet isocenter.”

**3. Exercise**

**3.4) (pg 7, line 271): consider replacing the idiom 'rule-of-thumb' with 'general guideline' (or something less colloquial)**

We have changed the line (line 316) to read: “NOTE: A **general guideline** is that an approximate 30% drop in PCr peak height corresponds to a Pi peak that is 50% of the height of the PCr peak.”

**4. Scan Protocol**

**4.7.2 (line 327): the note to record the start and end of each exercise interval is important; should this note appear earlier in the protocol, e.g., beginning of 4.7?**

We have moved this note to the beginning of the section (line 370).

**-4.11 Saving Data**

**4.11.2 (line 369): "If using a scanner" ? Is this in reference to a specific scanner/vendor?**

We have indicated that these instructions are specific for a Siemens scanner. “If using a Siemens scanner, select all spectroscopy acquisitions in the Navigator window.”(line 426, step 4.11.2)

**-Results and Fig 6B: reproducibility; the authors indicate a mean difference and stdev of 1.03 +/- 4.83s. I presume this refers to the mean difference and stdev of the recovery time constant, but this should be stated explicitly in the Results section (pg 10, line 406) and in the caption for figure 6B.**

The line (line 466) in the results section has been modified and now reads: “Bland-Altman analysis of PCr recovery time demonstrates a mean difference±standard deviation of 1.03±4.83s and between trials coefficient of variation of 4.66.”

The caption for 6b has been modified: “B) Bland-Altman analysis of 31PMRS technique reproducibility demonstrates mean difference±standard deviation of 1.03±4.83s for the PCr recovery time between trials the coefficient of variation is 4.66. ”

**-Results, pg 10, lines 417-418: "Additionally, note oscillations around the fit that are caused by the progressive loss of muscle control characteristic of this disease." Does this comment refer to the sinusoidal pattern in the recovery curve depicted in Figure 7? If so, please move the reference to Fig 7 from line 420 to the end of this sentence on line 418.**

Yes, this refers to the sinusoidal pattern and the reference has been moved as suggested. (line 481)

**-With respect to the oscillation observed in the participants with ataxia, is there any advantage to removing the oscillation (detrend) prior to fitting the curve to obtain the recovery time constant?**

This is an interesting consideration that warrants evaluation in a larger cohort of patients with ataxia.

**-Results, pg 10: very minor style suggestion: the latter part of the Feasibility Study section (lines 434-437) would benefit from the addition of articles ("the"): e.g. "The PCr recovery time constant..." instead of "Time constant of PCr recovery..." "The percentage of intramuscular fat..." instead of "Intramuscular fat percentage..." etc.**

All of the suggested modifications have been made.

***Additional Comments to Authors:* N/A**

**Reviewer #2:**

***Manuscript Summary:* The manuscript describes a simple method for estimating quadriceps 31P MRS PCr recovery kinetics following exercise as an estimator of skeletal muscle mitochondrial oxidative capacity (mOXPHOS). This method is relative old (~ 30 years) but is not widely used, largely due to the availability of 31P capable MRS scanners. The manuscript describes the technique adequately. However, the limitations and assumptions inherent in inferring mOXPHOS capacity from PCr kinetics are not well described. Simple reproducibility studies are provided for healthy participants as well as some comparative data between patients and controls**

***Major Concerns:* Overall more emphasis needs to be placed on how this method can be quality controlled. For example, how accuracy can be known when cellular PO2 is not measured, or how between site precision could be controlled or a multisite clinical study.**

**1. I disagree that 31P MRS measurement provided is a "direct measure of OXPHOS capacity". It measures PCr resynthesis following exercise which contains at least two assumptions: 1) PCr is resnythesized through the ATP hydrolysis by the creatine kinase reaction using ATP solely provided by oxidative phosphorylation; and 2) that the P/O2 ratio is constant. While in all likelihood these assumptions are met under the conditions of the experiment; the method is not direct. A direct method would assess intramuscular O2 consumption recovery e.g. as seen measured in single fibers by Wust et al., J Physiol. 2013 Feb 1;591(3):731-44; or possibly by near-infrared spectroscopy e.g. Ryan et al. J Physiol. 2014 Aug 1;592(15):3231-41. These papers should be discussed and cited in the main text.**

31P MRS is described as ‘direct’ in this manuscript to differentiate this technique from clinical functional measures, like VO2 measurements, which indirectly measure mitochondrial activity. However, the authors agree with the point that 31P MRS does not measure oxidative phosphorylation directly, and have modified line 74 appropriately: ” 31PMRS produces a **noninvasive**, *in vivo* measure of OXPHOS capacity in human skeletal muscle, rather than alternative measures obtained from explanted and potentially altered mitochondria via muscle biopsy, and relies upon only modest additional instrumentation beyond what is already in place on magnetic resonance scanners available for clinical and translational research at most institutions.”

The goal of this manuscript is to thoroughly describe a 31P MRS technique that can be used *in vivo* in humans with illustration of potential applications of this technique. We have added mention of the two other methods in the main text under ‘Significance of the technique with respect to existing/alternative methods’, noting potential limitations of the single fiber assay (invasive with attendant risks that make it less appealing for investigations requiring serial assessment) and NIRS-based approaches (penetration depth may be limited, particularly in obese patients where as little as 5 mm of fat attenuates the NIRS signal by 20% (Hamaoka T *et al*, Philos Trans A Math Phys Eng Sci. 2011), and the technique does not lend itself to the multidimensional assessment of muscle and other systems afforded by MR-based techniques).

**2. An additional crucial assumption, which is not addressed, is that mitochondrial O2 delivery is not limiting to OXPHOS and therefore to PCr recovery. It has been shown using lower body positive pressure and hypoxic gas breathing that the PCr recovery rate constant is sensitive to O2 delivery: PCr recovery after exercise is slowed by hypoxia or reduced blood flow. The recovery kinetics are also complicated during a metabolic acidosis. These need to be mentioned and any method needs to be able to account for or monitor adequate o2 supply.**

We have designed this exercise protocol to avoid metabolic acidosis, and have provided methods to ensure appropriate pH throughout the study. Impaired oxygen delivery could certainly impact the measurement; we have added to the text the exclusion of patients with “conditions that restrict blood flow or oxygen delivery to the lower limbs (e.g. peripheral artery disease)” (line 578) that our group and others use to minimize any confounding effect of impaired O2 delivery. The limitations section also emphasizes that the PCr recovery rate constant is sensitive to O2 delivery “Thus, in situations where the status of the microvasculature is in question, due to reduced oxygen supply or other factors, it would not provide an unambiguous indicator of mitochondrial status but rather represent an indicator of the *in vivo* status of the maximum oxidative ATP synthesis of muscle that may reflect some combination of OXPHOS and microvascular issues.” (lines 618-622)

**3. The sentence starting on line 180 is underdeveloped. You need to discuss: 1) by how much PCr flux is "dominated" by OXPHOS (not lactate-associated ATP production or other non-oxidative phosphorylative process), and 2) why PCr recovery kinetics represent OXPHOS capacity. This latter assumption only holds true for a first-order rate reaction (Michaelis-Menten reaction). OXPHOS control in muscle appears to be dominated in recovery by M-M first order kinetics (e.g. see Korzeniewski and Rossiter J Physiol. 2015 Dec 15;593(24):5255-68.), therefore the assumption is valid, but this needs to be proposed and discussed.**

We have cited the suggested work amongst additional references (line 207) that support: 1) a major component of PCr recovery can be attributed to OXPHOS and a smaller component (7-10%) from other sources, 2) PCr recovery can be modeled using a monoexponential model, and 3) OXPHOS control in muscle is dominated by the first-order Michaelis-Menten reaction.

**4. mOXPHOS is actually linearly related to the rate constant (k) of PCr recovery not the time constant (line 182). k=1/tau. Therefore OXPHOS capacity is, in theory, hyperbolically-related to tau of VO2 (and by implication tau PCr). This theory was proven in single cells in the presence of unlimiting O2 by Wust et al., J Physiol. 2013 Feb 1;591(3):731-44. 5. Line 202 and onwards. There is little discussion of alternative methods such as biopsy and NIRS assessments. Please highlight the advantages/disadvantages of each method.**

With the preponderant use of the time constant (tau) in the literature on *in vivo* muscle PMRS, we respectfully adhere to the use of tau in describing PCr recovery rather than switch to the rate constant. The specifics of the model are discussed later in the text (line 447). These are excellent references that we have now incorporated into the text (lines 636-647).

***Minor Concerns:* Line 74 see major comment 1**

We have expanded the discussion on alternative techniques as described under the response to major comment 1.

**88 specialized equipment.**

Line 88 (now line 86) is: “The goal of this work is to outline a reproducible method to measure *in vivo* skeletal muscle mitochondrial function in individuals possessing a wide range of abilities.”

**Mitochondrial impairment is 89 Friedrich's Ataxia is a bizarre condition to highlight, without mentioning far more common conditions abnormality in muscle OXPHOS capacity such as chronic heart, lung or kidney diseases, cachexia, mitochondrial myopathies etc. How about simple heathy aging?**

The overall goal of this manuscript is to describe a method with many potential applications, some more common and of greater interest to certain investigators than others. We along with the patients with FA whom we serve take issue with its characterization as a ‘bizarre condition’. Rather, FA as a condition whose distinct mitochondrial defect produces phenotypic changes similar to a broad range of disorders (e.g. diastolic heart failure, diabetes) makes it extremely relevant to this work. We have also highlighted other, more common disorders (i.e. aging, diabetes) as suggested (lines 102-123).

**104 How then do you explain the common finding in animal models that OXPHOS capacity is increased on exposure to a high fat diet? E.g. Jian et al. Diabetes. 2014 Jun;63(6):1907-13.**

The intent of line 104 (now line 112) is to describe a potential reason for reduced OXPHOS in the setting of lipid overload. There are instances of exceptions, and paradoxical findings (i.e. the athletes paradox, or the study referenced above). These interesting phenomena demonstrate the need for more accurate *in vivo* techniques with which to better understand instances where obesity does vs. does not associate with increased risk of cardiovascular disease.

**105 I don't agree with the concept of "number" of mitochondria. Mitochondria are a reticular conjoined structure. Their number is a meaningless construct.**

We have modified line 106 (now line 115) to read “This reduction in mitochondrial skeletal muscle OXPHOS capacity in the setting of lipid overload is accompanied by a decline in the quantitative (content and biogenesis of mitochondria) and qualitative…”

**113 changes in mitochondrial oxidative capacity are**

Line 113 is “However, it remains unclear whether **changes in mitochondrial capacity are** a cause or a consequence of insulin resistance.”

**124 cardiovascular capacity for what? Oxygen transport and utilization? Blood flow? What? Please be specific.**

We have changed ‘cardiovascular capacity’ to ‘cardiac output’.

**155-157 I am worried here that you are mixing the notions of cardiac mOXPOHS and skeletal muscle mOXPHOS. These two require very different approaches for measurement.**

The authors understand and acknowledge that the particular technique described here cannot be used for evaluation of OXPHOS in cardiac mitochondria. However, “Measurement of skeletal muscle OXPHOS capacity as detailed herein is readily implemented and robust; **coupled with the significance of skeletal muscle OXPHOS in heart failure**, these features **makes it an appealing biomarker in comprehensive studies of heart disease**[**36**](#_ENREF_36).” **(line 161)**The authors believe that this is a crucial, and often overlooked point prompting emphasis in the manuscript.

**181 see major comment 3**

We have cited additional references demonstrating that 1) a major component of PCr recovery can be attributed to OXPHOS and a smaller component (7-10%) from other sources and 2) PCr recovery can be modeled using a monoexponential model. (line 200-207)

**192 impaired mitochondrial oxidative capacity in. [N.B. mitochondrial have several roles in the cell; you need to say which capacity you are talking about. You could be talking about calcium retention capacity for example].**

**We have clarified this line: “**For instance, these techniques have shown impaired **OXPHOS** in subjects with type 2 diabetes.**” (line 217)**

**234 Why left leg? Why not dominant or non-dominant leg in all participants?**

While our setup uses the left leg, we have removed specification of ‘left’ when mentioning the leg in this protocol.

**271. see major comment 2. This type of severe intensity ballistic exercise can drive down PO2 very rapidly and therefore may limit PO2 availability for PCr resynthesis during exercise recovery. This may provide an artifact for extrapolation to OXPHOS capacity. Simply requiring a 30% drop in PCr peak height may still be accompanied by O2 delivery limitation.**

We have added mention of excluding patients with: “2) conditions that restrict blood flow or oxygen delivery to the lower limbs (i.e. peripheral artery disease)” (line 578) to minimize any confounding effect of impaired O2 delivery. Our lab and others have found that intact blood flow and avoidance of acidosis should prevent O2 delivery limitation, though adding perfusion imaging is certainly a technique that can be added to the MR protocol if this is a concern for a given study population.

**315 An improved protocol would acquire unsaturated GREs with full relaxation to first determine resting PCr peak area with unstaturated acquisitions. This allows the assumption of PCr;bATP peak ratio area to be applied (often 8.2 mM bATP is used) and estimate quantitative resting and exercise PCr concentrations.**

Thank you for this suggestion. We will explore this approach in future work.

**334 6 averages is not a whole number fraction of 20 seconds. How can the acquisitions during exercise be controlled so that the exercise is terminated exactly at the point at which one set of 6 acquisitions is complete and the other about to start. Not clear.**

We use 4 averages; the 6 was a typographic error.

**340 PCr depletion is not a measure of quality control. Signal to noise (variability) or precision (versus some known value) would be a suitable measure of quality control. The 'quality' of the protocol should be judged on the basis of some quantitative indicators of precision and accuracy of the method.**

This step refers to the quality of the exercise itself, not spectral or data quality. Steps to ensure quality of the overall acquisition are described in step 4.9.2 **“Verify that PCr peak height is the same at beginning of rest and end of recovery (typically <10% difference is desired).** This ensures that there was negligible loss of field homogeneity during acquisition.**”**

**391 why not fit this equation with k in the exponent? k is directly linearly proportional to OXPHOS capacity.**

We use tau given the preponderance of literature using the time constant rather than k.

**403 where these studies repeated on different days or the same day. Different days would be preferable.**

Studies were repeated on different days (within 1 week); this has been added to line 456.

**409 this only described reproducibility in healthy volunteers. In patients with O2 delivery constraints this protocol may be invalid for reasons discussed above.**

We have included an additional exclusion criterion for “2) conditions that restrict blood flow or oxygen delivery to the lower limbs (i.e. peripheral artery disease)” (line 578).

**518 the rationale for this 30% depletion target is not explained.**

One rationale for this target depletion is that reliable fits of the exponential recovery require sufficient dynamic range of PCr loss during exercise. We have empirically determined that a general guideline of ~30% depletion is sufficient for this purpose on our typical signal-to-noise ratio.

**542 mitochondrial number is not a valid variable. Volume-density perhaps? 548 ?**

We have modified line 106 to read: “This reduction in mitochondrial skeletal muscle OXPHOS capacity in the setting of lipid overload is accompanied by a decline in the quantitative (content and biogenesis of mitochondria) and qualitative…”

**560 see major point 1. Discuss NIRS**

Done.

***Additional Comments to Authors:* N/A**

**Reviewer #3: *Manuscript Summary:* This manuscript from Kumar and colleagues describes the use phosphorus magnetic resonance spectroscopy to assess skeletal muscle oxidative capacity in vivo in human participants. 31P-MRS has been extensively used in both rodents and humans to assess skeletal muscle metabolism, and thus this experimental video will provide a useful tool for the rookie MRS user, although instrumental differences between magnets will make exact replication more difficult.**

**Introduction -Line 86-88: Can you really claim this this technique doesn't require specialized equipment? It is required that you have a magnet capable of multi-nuclear spectroscopy (not always the case especially in clinical situations) and a tuned coil for phosphorous detection.**

In this case, specialized equipment refers to MR-compatible ergometers and other additional equipment beyond the necessary materials for phosphorous spectroscopy. The authors understand that access to MR and spectroscopy equipment is limited, and we have deleted this portion of the sentence.

**-Line 93-113: While mitochondria have been associated with many disease states, impaired maximal capacity for oxidative phosphorylation is rarely the cause for disease. This is particularly evident in the case of insulin resistance/T2D where oxidative capacity is not impaired during the development of insulin resistance, but appears only with chronic T2D, suggesting this is not causative (See work from Fisher-Wellman et al. Diabetes 2014). Furthermore, the description of decreased oxphos capacity causing lipid accumulation is misleading and not well supported by the literature. Most cells, particularly skeletal myocytes, normally function with a basal energy demand which is far lower than the oxidative capacity. It is not clear how the authors link modest reductions in the maximal capacity associated with T2D would result in lipid accumulation. Wouldn't there have to be alterations in the basal 'energy balance' for this to occur? Clearly this is not the point of this methods paper, but these topics are often misconstrued in the literature.**

The use of this technique is more to observe changes in mitochondrial function in the setting of T2D and other diseases, rather than determining a mechanism for these changes. We agree that statements regarding causation could be potentially misleading and have taken care to avoid this in the manuscript.

**Line 130-149: It is unclear why there is a paragraph discussing mitochondria and cardiac dysfunction since the methodology described cannot be used to assess cardiac mitochondria function. Granted there are capabilities for 31P-MRS in cardiac metabolism, the exercise/recovery protocol is not possible. It is suggested that this paragraph be removed.**

The authors understand and acknowledge that the particular technique described here cannot be used for evaluation of OXPHOS in cardiac mitochondria. However, “Measurement of skeletal muscle OXPHOS capacity as detailed herein is readily implemented and robust; **coupled with the significance of skeletal muscle OXPHOS in HF**, these features **makes it an appealing biomarker in comprehensive studies of heart disease**.” (line 161-164)The authors believe that this is an often-overlooked point.

**Line 172: phosphorus in ATP does not contain 'high energy'. This is an enormous misconception in bioenergetics. In bioenergetics, the potential energy for cellular work comes from the extent to which the observed mass action ratio is displaced from equilibrium (i.e. ΔGATP), which in the case of skeletal myocyte cytoplasm is ~-60kJ/mol. This is where the driving force for cellular work is derived, not the 'phosphate bond'.**

We describe phosphorous as a component of high-energy substrates. The line does not imply that phosphorous itself contains the energy to do useful, cellular work. Line 172: “Magnetic resonance spectroscopy tuned to **phosphorous (31PMRS), an endogenous nucleus found in various high-energy substrates** within cells throughout the body”. (line 186)

**Line 180: The authors should more clearly describe the energetic model used in this protocol. Upon cessation of exercise, ATP production is primarily driven by mitochondrial oxidative phosphorylation. This ATP is 'shuttled' from the mitochondrial matrix to the cytoplasm where is can be utilized by CK to resynthesize PCr until the cellular equilibrium for PCr/tCr is reached (resting). The authors should also mention the potential for other sources to contribute ATP during the recovery phase (see work from Sean Forbes and Ron Meyer which shows ~7-10% contribution from glycolytic sources).**

Thank you for these excellent references. We have referenced the work of Forbes and Meyer in lines 202 & 207.

**Line 185: Several pivotal validation studies are not cited including: Meyer (1988) Am J Phyiol McCully (1993) J Appl Physiol Larson-Meyer (2000 ) Muscle and Nerve Lanza (2011) J Magnetic Resonance Imaging Layac (2016) AJP Endocrinology and Metabolism**

We have included references to these important validation studies in line 207.

**Line 182-184: There are several approaches for modeling this approach including the very simple calculation of a time constant. It is recommended that the authors provide a discussion regarding the various analytical approached and briefly the benefits and limitations. See elegant work from Gwenael Layec and Russ Richardson.**

We have included references to these studies in the manuscript. Further discussion of the model and calculation of the time constant has been included in the methods step 5.4.

**Protocol**

**General: It is suggested that the authors provide some general background regarding the safety of MR testing. Many safety precautions are taken to ensure that participants can have MRI testing and all equipment used and personnel involved in testing are cleared. Clearly the testing will occur under the supervision of an MR technician, but I think it is critically important to provide a little background on this given its importance.**

Thank you for this important point. We have added this at line 296-297: “It is critically important that all procedures involving MR equipment must be performed by adequately trained personnel adhering to the highest standards of MR safety,” and cited MR safety guidelines.

**Line 239: These instructions should be described for a surface coil. Volume coils will by slightly different. It should also be noted that testing of other muscle (besides the VL) is possible depending on the positioning of the surface coil.**

The instructions in this manuscript are specific to the equipment indicated (i.e. 10 cm surface coil). The exercise method described in this manuscript is specific to the VL and therefore we did not include a discussion of testing other muscles. Testing of different muscle groups could be an interesting follow-up to this publication.

**Line 243: Unsure what 'baby oil' is used for? Localization can be easily performed with scout images…**

The baby oil is used to mark the position of the coil for localization on the proton image. This position is used to guide the selection of the shim volume for B0 field homogenization. Some phosphorous coils come equipped with a fiducial marker. This is a simple alternative approach for coils that do not.

**Line 277: Actually, using the recovery kinetics approach for PCr, it is not necessary to account for differences in exercise duration because the mono-exponential model is used for fitting (i.e. first-order metabolic system). The only aspect that needs checking is to see if pH has changed too much…**

The duration of exercise is important to this technique because it informs us of the exact end of exercise PCr measurement, which is crucial for analysis.

**-Line 288: These sequence parameters will be specific for their equipment. This should be noted here for readers/viewers. It is also possible to measure PCr kinetics with faster time resolution which might be required for those participants with faster PCr recovery.**

We have generalized the descriptions in this section to ensure that any specific values can be replicated on different systems (even if the technique to set those values may be different). We have also ensured that any system specific instructions have been noted as such in the text.

**Line 293: The authors should be clear here regarding shimming. Typically shimming occurs using the proton coil and signal from water within cells. Moreover, the practice of shimming involved more than moving the box, it involves small changes to the magnetic field coils to make the signal most homogenous.**

We fully appreciate that the physics of proton shimming involve more than moving a box. For this practical methods paper, we have detailed the steps necessary to acquire good quality phosphorous spectra, and have omitted ,due to scope, a discussion regarding the physics of proton shimming.

**General comment: The authors suggest making several scans in which 20 spectra are acquired. In most cases, this protocol is run using a single scan with multiple spectra acquisitions (for example acquiring spectra every 6 sec for 5 min). Having to start new scans during and after exercise will result in a loss of some signal as PCr is being resynthesized while the next scan protocol is sent to the scanner…. Moreover, it is important to recognize that the number of scans required for accurate exponential fitting will depend on the recovery time constant. For example, the slower the recovery the more scans will be required to properly and accurately fit the data mathematically to an exponential.**

The method described in the protocol ensures **“that post-exercise acquisitions begin immediately following the exercise sequence, without pause or shimming”** (step 4.8.1). Splitting of the recovery scans is to allow for dynamic observation of depletion and recovery, which is only possible after acquisition is complete on our particular system.

**Line 334: I would disagree with the author's comment regarding splitting the recovery if exercise needs to be repeated. It is vitally important that enough time is given for the muscle to fully recover from exercise as calcium and other ions can alter mitochondrial bioenergetics and bias the results. Typically, we allow for a 10min rest period between exercise/recovery measurements to ensure this does not occur.**

We completely agree that sufficient recovery is important if repeating the exercise after insufficient effort. In light of inescapable time constraints on *in vivo* MR experiments, particularly in human subjects, our proposed scheme allows us to interleave other scans in the exam during the recovery before the next exercise bout. We have now clarified this in the text.

**-Line 358: The authors should note why it is important to measure pH. Tell the readers what impact this can have on the results…**

We have included a brief note on the significance of pH in OXPHOS recovery, along with a reference to the work by Jubrias *et al.* J Physiol 2003. (line 653)

**Line 384: The AMARES algorithm does not calculate Peak heights, but rather calculated the 'area under the peak', which is more accurate for calculating absolute concentrations of metabolites.**

We thank the reviewer for highlighting this confusing nomenclature of the jMRUI software. We have clarified in the manuscript that the AMARES peak area, referred to as “amplitude”, is used in the analysis.

**General comment: The authors do not describe any quality control measures taken to assure curve fitting is of high quality. The software will always force a curve fit and produce a time constant, but this may not be an accurate representation of the muscle oxidative capacity. This is an important point of discussion especially considering the large introduction suggesting the use of this approach for testing the efficacy of therapeutics…**

We have included step 5.7 that describes a method to obtain goodness-of-fit measures.

**Line 555: Varying levels of PCr depletion do NOT impact to the recovery of PCr. In fact, the main benefit of this modeling approach is the recovery kinetics (1st order exponential) are completely independent of the amount of PCr depletion. (see work from Ron Meyer 1988 and 1989 Am J Physiol).**

We thank the reviewer for helping us clarify this point. We did not intend to suggest recovery kinetics were directly related to the degree of depletion but rather that the fitting of noisy data to an exponential recovery curve can be impacted by this and so sufficient dynamic range of the signal is important in the realm of limited signal-to-noise. We have now clarified this in the text.

**-General comment: The major limitation of 31P-MRS is that if a defect in oxphos is indicated, this approach does not provide any mechanistic insight. The same goes for potential therapies. Its major benefit is the easy and repeatability of measuring human participants.**

We thank the reviewer for providing these comments. We have emphasized the idea that this technique may not provide specific mechanistic insight into the development of disease, but rather offers a useful biomarker of disease status. “It is important to note that unlike muscle biopsy analysis, which permits the measure of specific mitochondrial characteristics such as mitochondrial content and size as well as the mitochondrial maximum ATP synthesis rate, the *in vivo* measure using 31PMRS represents an aggregate of these direct measures in addition to extra-mitochondrial factors such as microvascular supply of oxygenated blood to the muscle.” (line 657-661)

**Figures**

**Figure 1: Not sure this figure is all that relevant to this approach. Small musculature exercise does not result in the same systemic effects as whole body exercise. I think it would be better to provide a metabolic model for the PCr model used within. This will make is easier for readers/viewers to understand how this measurement works and the assumptions made regarding the bioenergetic system.**

Figure 1 helps put these methods into what we feel is highly relevant context for investigators seeking a technique to better understand integrative physiology *in vivo* in humans with a broad spectrum of cardiovascular and metabolic disorders.

***Major Concerns:* N/A**

***Minor Concerns:* N/A**

***Additional Comments to Authors:* N/A**