Author’s response to reviews

Title: ROCKETSHIP: a flexible and modular software tool for the planning, processing and analysis of dynamic MRI studies

Authors:

   Samuel R Barnes (srbarnes@caltech.edu)
   Thomas SC Ng (thomas.ng@uci.edu)
   Naomi Santa-Maria (naomism@caltech.edu)
   Axel Montagne (montagne@usc.edu)
   Berislav V Zlokovic (zlokovic@usc.edu)
   Russell E Jacobs (rjacob@caltech.edu)

Version: 6
Date: 26 April 2015

Author’s response to reviews: see over
Dear Dr. Tan and Mr. Gato,

RE: MS# 2897421615575663

We would like to thank you and the reviewers for providing informative and insightful feedback about our manuscript.

All comments offered by the reviewers have been addressed and the required modifications/updates to the manuscript made. Specific details are discussed after each comment (attached) and are highlighted in red.

Please do not hesitate to contact me if any questions or comments arise.

Thank you for your consideration. We look forward to your correspondence.

Yours Sincerely,

THOMAS NG, M.D., PH.D.
BECKMAN INSTITUTE
DIVISION OF BIOLOGY AND BIOLOGICAL ENGINEERING
MAIL CODE 114-96
FACSIMILE (626) 449-5163
TELEPHONE (626) 395-5833
EMAIL THOMAS.NG@UCI.EDU
1. The authors present a new software package for DCE-MRI analysis, and motivate this by the observation that all existing packages are "limited" in one way or another (p5, 129). This is undeniably true, but to address this problem by creating a new package is but one solution. The alternative would be to join an existing project (quite a few of them are open source) and help to fix the limitations. In the long run, I believe that the latter is a more constructive solution. The creation of new packages leads to a fragmentation in the field that actually hinders rather than promotes a wider use of these techniques. The authors should provide a stronger motivation for their chosen approach.

The development of ROCKETSHIP was motivated by the diverse needs of our collaborators and members of our research group. The skill sets within this community ranged from experienced programmers to biologists/clinicians with little coding experience. Since each individual often has unique application requirements of the software, it was important that the software package can be used, developed and extended by all these people. Matlab was thus chosen because it can be accessed at multiple levels of expertise. It has a strong built-in and user-contributed library of image processing and statistical functional capability while retaining features that allow portability with lower-level languages. Furthermore, the interpretive nature of Matlab also facilitates module development at all levels of coding expertise. Finally, Matlab is available at most academic institutions, enabling widespread package adoption and use.

We acknowledge that an alternative approach is to extend an existing project. However, given our past experiences with available software, as well as the needs of the users that drove the development of ROCKETSHIP, we felt that the development of this software package using Matlab was appropriate and complements the offerings currently available. We have strived to ensure that the data inputs and outputs of ROCKETSHIP are compatible with other software, enabling portability between software suites if required.

We have updated the manuscript to highlight our rationale for developing ROCKETSHIP. (Lines 129-149, 351-359)

2. This will also require a more in-depth review of existing packages, as well as their possibilities and limitations, to identify exactly where ROCKETSHIP proposes to present an improvement. The list provided by the authors (p5) is incomplete, eg. packages such as dcemri (http://dcemri.sourceforge.net/), dcetool (http://thedcetool.com/), or pmi (https://sites.google.com/site/plaresmedima/) are not included. Considering this more extended list, I am not convinced that the limitations pointed out by the authors are not already addressed. For instance, I’m not sure it’s justifiable to say that dcemri does not enable data-driven methods.

In addition, though this is less relevant for the publication, the list of commercial software packages is incomplete. In order to avoid the appearance of bias, the authors should either present a complete list of commercial packages, or else leave them unnamed.
We have added a table (Table 1) summarizing the capabilities of currently available software packages. Further rationale for ROCKETSHIP is addressed above and in the manuscript. (Lines 351-359)

While dcemri can estimate parameters using Bayesian estimation, recent data-based methods such as nested model selection are not implemented in dcemri (or currently available software). We have clarified our use of the term “data-driven” to reflect this (Lines 133-135).

We have removed the listing of commercial packages.

3. Before the implementations section, it would be good to include an explicit list of the product specifications. What exactly is the functionality? Is it designed to be modular/extendable? What is the target audience for this application (eg. radiologists? Physicists? Biologists? Cardiologists? …), what level of training or expertise is expected, what level of support will be available, product lifetime etc..

We have expanded the introduction to clarify the functionality of the software suite. (Line 140-149).

The target audience include radiologists, biologists, physicists and engineers performing preclinical and clinical work with at least moderate computer expertise and some DCE-MRI experience. For that audience, we expect that minimal training will be required, although the authors will be available via email. The lifetime of the current instantiation is expected to be 3-5 years, but given the extensibility we anticipate updates will be made for individual purposes, with significant updates at least on a yearly basis; updates will be made available on github. We have updated the discussion to reflect this. (Lines 400-409)

4. Software validation (p11-13): it seems to me this paper has a dual purpose: on the one hand, a novel core algorithm for DCE-MRI is proposed (p8-9), and this is validated with application to simulations and two datasets; one the other hand the paper aims to validate the new software designed around this algorithm. These are two entirely different objectives, and I don’t think they should be combined into one paper as each needs to be performed in considerably more depth in order to be convincing. A novel fitting algorithm can only be validated by a much broader simulation study and by explicit comparison against existing algorithms in a range of relevant tissue types – this to show that it represents an improvement compared to existing state-of-the-art. This should be done outside the context of a specific software for the sake of transparency and generality. It also requires a higher level of detail in the description of the methods.

The nested model algorithm implemented in ROCKETSHIP was described and examined in prior studies by Ewing et al [1-3]. While these studies provided validation of its use for varying tissue permeabilities, a useable software implementation for the general public is currently unavailable. Thus, we felt it appropriate to include this model option in ROCKETSHIP.

To ensure that this implementation of the algorithm performs as described in prior studies, data simulating the different levels of nesting and SNR was performed demonstrate its ability to recover the appropriate nesting level at different SNR (Figure 5), which can affect the fitting accuracy of all models [2].
As discussed in the manuscript (Lines 393-398) and in prior literature [2], the optimal selection of model/algorithm will be dependent on several factors, ranging from acquisition parameters to the specific research/clinical question at hand. Availability of a software suite that provides a range of models and algorithms while allowing easy modifications/extension to the software for individual purposes shall aid this effort for both entry-level and advanced users.

The validation of a software tool itself requires a test against the product specifications (which are too vague at this point – see 3). Here I would expect a comparison of results obtained with standard algorithms, to other open-source softwares that implement the same algorithm. This to check that implementation differences do not affect the result.

Validation of the software implemented algorithms in this study was performed by testing the ability of the software to recover DCE-MRI parameters generated via simulation. The reliability and accuracy of the software algorithms across a range of conditions were tested by performing Monte-Carlo simulations at different SNR, time resolution across a range of DCE-MRI parameters and models. Prior studies have adopted this standard (using the QIBA simulated data), to validate the ability of a software algorithm to recover appropriate DCE-MRI parameters [4, 5]. We did not use the QIBA dataset in the current study since it was only limited to the Tofts model and not applicable for the majority of the models implemented. However parameter recovery using ROCKETSHP compared favorably to studies using the QIBA dataset. Discussion regarding this has been added to the manuscript (Lines 269-280).

I would expect at this point also more extensive testing with specified users. If, for instance, the software is designed to be suitable for inexperienced radiologists, the results produced by such users could be compared against those of a collection of experts. As for the algorithm, this could also involve a comparison against other softwares in the same users to check whether ROCKETSHP offers an improvement in that sense. Does the new software allow users to get their results more quickly? With less training?

We agree that the human-computer interface aspect of radiology software is an interesting area of study and probably requires more exploration. However, such a study is beyond the scope of the current manuscript. The current manuscript introduces a DCE-MRI data processing environment using MATLAB, and validates the ability of algorithms implemented within the software to recover DCE-MRI parameters appropriately. We hope that implementation of this software in a widely used computing environment will facilitate DCE-MRI analysis for both novice and experienced users. We are not aware of any publically available software that currently implements the nest-model algorithm.

5. I don’t think the appendices are necessary in a publication. All of this material can be referenced.

We agree that the appendices can be shortened. However, there are details within the appendices, especially in Appendices B-D, that are pertinent to the implementation of ROCKETSHP. Availability of this information allows the user to make an informed comparison with prior efforts (which mostly provided similar information) and to modify the current package for their individual purposes.

We have streamlined the appendices to this effect.
1) The authors claim a superiority of their software in several terms including the number of models etc. However, named softwares in the introduction provide similar methods/models etc. Please tone down the argumentation or provide a head to head comparison to show how superior your approach is.

We have added Table 1 comparing the features of currently available software. The rationale for the development of ROCKETSHIP has been expanded within the manuscript (Lines 129-149, 348-356).

2) Please add PMI as another research software providing DCE MRI (https://sites.google.com/site/plaresmedima/)

PMI has been added in the introduction (Line 125, Table 1).

3) The authors describe on the one hand their software and also introduces an approach to employ nested models. To me both part are described not in detail and the novelty and uniqueness of these approaches is unclear.

For example, as the authors correctly mention, there is a broad body of such perfusion analysis software. Most of what is reported and implemented has been shown before. So what are the new, unique parts or is the implementation enhanced? I would suggest to split the manuscript and separate the nested modeling from the actual software implementation. By this the limited description on how exactly the parameters of the nested models are judged etc can be extended while in another paper the software can be outlined in more details on the implementation.

The nested model algorithm implemented in ROCKETSHIP was described and examined in prior studies by Ewing et al [1-3]. While these studies provided validation of its use for varying tissue permeabilities, a usable software implementation for the general public is currently unavailable. Thus, we felt it appropriate to include this model option in ROCKETSHIP.

To ensure that this implementation of the algorithm performs as described in prior studies, data simulating the different levels of nesting and SNR was performed to demonstrate its ability to recover the appropriate nesting level at different SNR (Figure 5), which can affect the fitting accuracy of all models [2].

The development of ROCKETSHIP was motivated by the diverse needs of our collaborators and members of our research group. While current software offerings collectively provide a wide range of DCE-MRI processing functionalities, no single software provided a streamlined pipeline to process datasets with access to multiple kinetic modeling options. Furthermore, extensibility of current software can be problematic for inexperienced users; software implemented in IDL and C need to be recompiled prior to further use. Matlab was thus chosen because it can be accessed at multiple levels of expertise. It has a strong built-in and user-contributed library of image processing and statistical functional capability while retaining features that allow portability with lower-level languages. Furthermore, the interpretive nature of Matlab also facilitates module development at all levels of coding expertise. Finally, Matlab is available at most academic institutions, enabling widespread package adoption and use.

As discussed in the manuscript (Lines 393-398) and in prior literature [2], the optimal selection of model/algorithm will be dependent on several factors, ranging from acquisition parameters to the
specific research/clinical question at hand. Availability of a software suite that provides a range of models and algorithms while allowing easy modifications/extension to the software for individual purposes shall aid this effort for both entry-level and advanced users.

We have modified the manuscript to emphasize this point (Lines 129-149, 351-359).

4) There several topics mentioned in this manuscript which does not contribute to the topic actually. For example what is the purpise of fitting the ADC when you just measure DCE-MRI ?? Furthermore, you describe in the appendix three different appraoches to measure and calculate T1 maps but actually you just used the VFA approach in this study ?
Please remove everything that does not belong to this study from the manuscript!

We agree that the T2/T2* and ADC mapping do not specifically pertain to the DCE-MRI aspect of the software. However, these two techniques are often used in conjunction with DCE-MRI studies and was implemented in ROCKETSHIP for our group’s use.

From our experience, the majority of current DCE-MRI studies (and the studies at our institution) use the VFA approach. However, other methods to derive T1 were included in the current implementation for our group’s use.

Given that users of ROCKETSHIP may find these functionalities useful, we have kept these functions in the current software and thus describe their implementation accordingly.

5) The evaluation on real world data is limited. Especially there was a single (!) ssample of preclinincal and clinical data used here. Even more, there is was no gold standard approach as comparison used. How do you know that the estimated parameters are correct when using a single case ? Please provide more data sets and a reference method for estimating the error in your results.

Accurate and reliable estimation of real-world DCE-MRI data continues to be a challenge. It is likely that several models are necessary to capture the range of permeability values observed in vivo. The ability of kinetic modelling to accurately estimate permeability values is one aspect of the puzzle. The goal of this manuscript is to introduce a software suite that can process and analyze DCE-MRI data using multiple contemporary kinetic models (including the nested-model) as well as statistical tools to compare the fitting of those models. Simulations provided a reliable reference standard to evaluate the precision and accuracy of the model fitting algorithms to recover the expected permeability values. Simulation data has been used in prior studies to examine the functioning characteristics of various pharmacokinetic models [6] and validation of prior software [4, 5].

The real world animal and human data provided demonstrated that the software can be reasonably applied to in vivo data.

6) There several flaws in the imaging of the clinincal case. To me this subject is a healthy patient assumable with no BBB leakage ? Given the imaging parameters any first pass of the CA is missed in these settings so I wonder how
the AIF was measured. Usually, for DCE and DSC temporal resolutions are much higher.
Secondly, usually images are take in the axial plane rather than coronal. All this needs much more explanations. Why not use a standard exam described in the literature before for the demonstration of the software?

The clinical study was a standard protocol performed at our institution as part of a larger study examining the BBB in healthy subjects and dementia patients. Results for this study have been reported in [7]. The AIF was derived from the internal carotid artery and shown in Figure 7. We have clarified the manuscript to convey this fact (Line 327).

We agree that several image acquisition parameters can affect accurate and precise estimation of permeability values. Exploration of this topic is one of the major motivations for the implementation of ROCKETSHIP. We are currently exploring this issue with further imaging studies and analysis with this software.

7) I do not see actually the benefit of Matlab, except for basic research questions where people want to test out new methods to analyse dce mri. The software is built on several add on toolboxes that have to be purchased onto of the Matlab license. Moreover, it seems that the software was not compiled for use outside the Matlab IDE. This makes handling this difficult in a clinical research setting (from my experiences).
Furthermore, as correctly stated by the authors, the processing in Matlab is slower since the instruction are interpreted rather than they are compiled. To process whole image maps or larger 3D datasets even with parallel imaging TB this is costly.
Did you try to do some smart implementations in matlab (linearization etc?).

We agree that Matlab would not be the programming environment of choice if one was to process data in a pure clinical workflow scenario. However, as discussed in the manuscript and by Heye et al [8], further preclinical and clinical research towards accurate estimation of tissue permeability is needed. From this standpoint, software that is accessible at all levels of programming expertise, from the engineer/physicist to the biologist/physician, would be ideal. Matlab was thus chosen because of its flexibility. It has a strong built-in and user-contributed library of image processing and statistical functional capability while retaining features that allow portability with lower-level languages.
Furthermore, the interpretive nature of Matlab also facilitates module development at all levels of coding expertise. While Matlab is a commercial package, it is available at most academic institutions, enabling widespread package adoption and use of the software package. We have updated the Discussion to convey this point (Lines 351-380).

Where possible, we have implemented linearizations of the code to improve throughput and make use of code parallelization that Matlab offers to accelerate the processing.

8) You just provided data for the patlan and tofts model, how does the other implemented models perform on your data?
Tables 3-5 have been updated to demonstrate the ability of all the implemented models to recover simulated data values.

The ability of the nested model algorithm to recover the appropriate simulated model is shown in Figure 5 and Table 6.

9) Figure 2 and 4 can be dropped since they did not contribute to the manuscript.

We have removed both Figures from the manuscript.

10) I miss an actually results section that clearly outline the performance of the software and the nested model analysis. This is somehow mixed with materials and methods and description of the software. Please provide a more clear structure of the manuscript.

We have modified the manuscript to delineate the sections describing the methods and corresponding results for regular model fitting and nested model analysis respectively. (Lines 255-290).

11) Please revise the discussion:
- head to head comparison of the different software available (just descriptive)
- more extensive discussion of the novelities and implementation compared to others - critical discussion of the limited number of data sets and missing refernce standard

We have updated the Discussion to address both comments (Lines 350-397).

12) Please abbreviate (or not) the journal titles consistently as provided by the journal's guidelines.

We have corrected the references to reflect the journal’s guidelines.
Bibliography & References Cited


