Suggestions for improving the reporting of magnetic resonance spectroscopy voxels and spectra

Vuong Truong^{1,2 †}, Niall W. Duncan^{1,2 † *}

1. Graduate Institute of Mind, Brain and Consciousness, Taipei Medical University, Taipei, Taiwan
2. Brain and Consciousness Research Centre, TMU-ShuangHo Hospital, New Taipei City, Taiwan

† Both authors contributed equally to the work

* Corresponding author

Graduate Institute of Mind, Brain and Consciousness
Taipei Medical University
172 Keelung Road, Section 2
Taipei, Taiwan
niall.w.duncan@gmail.com

Abstract

Magnetic resonance spectroscopy (MRS) has seen an increase in popularity as a method for studying the human brain. This approach is highly dependent on voxel localisation and spectral quality, knowledge of which are essential for judging the validity and robustness of any analysis. As such, visualisation plays a central role in appropriately communicating MRS studies. The quality of data visualisation has been shown to be poor in a number of biomedical fields and so we sought to appraise this in MRS papers. To do this we conducted an informal survey of the MRS literature, focusing on papers comparing people with psychiatric disorders to controls. Data visualisations in 50 papers were given scores for the amount of information conveyed about voxel localisation and spectral quality. This revealed a generally low standard, with a significant proportion of papers not providing the information required to judge the validity of their findings or replicate the experiment. Some basic suggestions are then presented as to a minimal standard for MRS data visualisation. Sample code to produce the suggested figures is provided.

Keywords: MRS; methodology; spectroscopy; MRI; data visualisation; robust

Introduction

Magnetic resonance spectroscopy (MRS) has seen an increase in popularity as a method for studying the human brain. The technique allows the *in vivo* estimation of metabolite concentrations within specific brain regions, opening up a variety of interesting research questions. For example, researchers have used MRS to draw connections between individual differences in psychological processes, such as working memory (Marsman et al., 2017; Takei et al., 2016) or inhibitory control (Cheng et al., 2017; Hermans et al., 2018), and neurotransmitter concentrations in a variety of brain regions (Duncan et al., 2014). This complements a longstanding use of MRS to study brain disorders. Here, metabolite levels have been compared between patients and controls, revealing neurochemical differences in conditions such as depression (Hasler et al., 2007), schizophrenia (O'Callaghan et al., 1991), autism (Minshew et al., 1993), ADHD (Maltezos et al., 2014).

In both these types of study - looking at regional correlates of individual differences or comparing different groups - the data being analysed must necessarily come from the same part of the brain. If they do not then any correlations become difficult to interpret and spurious differences between groups may arise through systematic differences in the location sampled. Studies have, however, demonstrated test-retest voxel overlaps of only around 65% for PRESS voxels (Lee et al., 2013) and 75% for MEGA-PRESS (Bai et al., 2017). This leaves scope for there to be relatively high location variability across participants and, by extension, for group differences in location to arise. Given this, it is important when publishing work for readers to be given information about the degree of overlap achieved in a study. It is not possible to judge the robustness of the work without this information, nor is it possible to identify any confounding factors arising from differences in voxel locations across participants.

A second key factor for interpreting MRS studies is the quality of the data included. Although summary statistics about the spectra themselves (e.g., signal-to-noise ratio) and the quality of the analysis fit (e.g., fit error) can provide a guide to this, the final judgement must generally involve visual inspection (Kreis, 2004). The information that this judgement relies upon should not, however, be available only to the researchers involved. As with voxel locations, it

is important that readers of published work also be able to directly judge the quality of the data included.

In both cases, the issue for readers is essentially one of data visualisation. An effective representation of voxel locations would include information about all participants included in the analysis, presented in such a way as to show localisation consistency. Where two or more groups are involved, an effective visualisation would also present information about group-specific locations. For MRS spectra, an appropriate visualisation would present the data of all participants. Group membership would be important additional information here too, to allow any patterns in data quality between groups to be identified. Providing all this information then allows the reader to appropriately interpret results and judge their reliability.

This need for effective data visualisation to increase the interpretability and robustness of results has been highlighted in other areas of science (Weissgerber et al., 2015). This includes the brain imaging field, where there have, for example, been calls for more informative data visualisation in EEG studies (Rousselet and Pernet, 2011), along with various discussions about how to represent fMRI data in a robust and informative manner (Lau et al., 2019; Stelzer et al., 2019). Unfortunately, the quality of data visualisation has been found to be low in a significant proportion of publications across the biomedical literature (Chen et al., 2017; Hertel, 2018; Weissgerber et al., 2019).

Literature appraisal - method

Given the low quality of data visualisation in other areas, to what extent do MRS publications currently provide the information described above as being necessary to appropriately interpret them? To gauge this, we conducted an informal appraisal of the recent literature and rated the data visualisations provided in 50 papers that compared individuals with mental or behavioural conditions to neurotypical groups and noted a range of reporting factors for each. Work that compared groups was chosen as MRS studies in these areas are relatively common and is potentially affected by issues of systemic location differences or spectral quality issues.

In more detail, we searched the Pubmed database with the following search term: “(MRS OR (magnetic resonance spectroscopy)) AND (depression OR depressive OR schizophrenia OR

schizophrenic OR bipolar OR hyperactivity OR ADHD OR autism OR autistic)”. Only original human research papers written in English and published in the past five years (from 2015/01/01 to 2019/12/31) were considered. This search produced 611 results. We then took the first 10 papers suggested by the Pubmed “Best Match” feature for each of the five conditions that reported single voxel MRS results.

Having selected a sample of 50 papers, we firstly appraised how well a reader could judge the consistency of voxel locations across participants and between groups. Visual representations of voxel locations, when present, were categorised according to whether the location was indicated using real data or whether an example location was drawn on manually. Ambiguous cases were marked as being from data (2/50 papers). Where real data was used, we further categorised the type of representation according to whether a single example participant was shown or whether data from the whole group was shown. The best case would be a figure created from all participant data that represented the degree of group overlap in some way.

To complement this information about visual representations, we also appraised textual descriptions of the methodology used to cite the MRS voxels. These were graded into three categories: Category 1 reporting consisted of a non-specific location such as the lobe or general brain region targeted (e.g., occipital lobe, DLPFC, etc). Category 2 reporting included a description of the specific region covered by the voxel (e.g., the voxel covered the pre- and post-central gyri). Finally, Category 3 reporting included the specific anatomical landmarks used for locating the voxel and how it related to these (e.g., “the anterior edge was aligned with the genu of the corpus callosum”). How the voxel size was reported was also noted, along with whether any location coordinates or information about voxel overlap was given. Ideally, papers would provide a Category 3 methodological description, along with the voxel dimensions.

The second area that we appraised was how well a reader could judge the spectral quality for the data included in the study. For this, the visual representations of MRS spectra (where present) were categorised according to whether there was an example spectrum from a single participant; a figure showing the average spectra of a group or groups; or a figure presenting

the spectra from all participants. The latter visualisation would be the optimal way of representing the data.

Literature appraisal - Results

Voxel location

The majority of studies provided a figure representing the voxel location but 14% did not. For those that did provide a figure, the most common approach was to create an illustration of the location by drawing over an anatomical image (53%). This was closely followed by showing the true location for one participant (i.e., a voxel mask created from the data) superimposed upon their anatomical image (47%). In terms of representing the variability of voxel locations between participants, very few studies showed the location of more than one person: One study showed the location for each participant and one study provided a representation of the group overlap.

In terms of the methods text, the largest proportion of studies provided a Category 1 description of the voxel location (48%). In other words, their methods only related a non-specific anatomical location with no details of landmarks used during scanning. The second most common type of description was then Category 3 (30%), meaning a specific location was provided, often along with a description of the anatomical landmarks used to locate the voxel during scanning. The remaining studies gave a Category 2 description (22%). When reporting the size of the voxel, the most common approach was to report the dimensions of each side (88%). The remaining studies reported the voxel volume (4%) or, in some cases, gave no information at all (8%). No studies reported the degree to which the MRS voxels for different participants overlapped with each other, nor did any studies that compared different groups report how similar the voxel locations for each group were to each other.

Visual representation of spectra

A quarter of studies did not provide any visual representation of the MRS spectra (26%). Of those that did, the majority showed an example spectrum from a single participant (68%), along with a small number that showed a group average (6%). Only two studies provided a visual representation of the spectra for all participants. These two papers were also the only

ones to provide figures that allowed a comparison of the spectra between the groups as a whole. Notably, the studies that presented the data from a single participant did not provide any rationale for why that participant was selected.

Suggestions for reporting

From this appraisal of the literature we can see that there is a generally poor standard of reporting of voxel locations and spectral quality. This low standard in many papers makes proper interpretation of the results or replication of the studies difficult or, in some cases, impossible. To improve this situation we suggest a minimal standard for the visualisation of MRS voxels and spectra. This is intended to complement other work discussing optimal approaches to MRS data acquisition and analysis (Mullins et al., 2014; Wilson et al., 2019), as well as a recently proposed approach to assessing the overall quality of reported MRS studies (Peek et al., 2020). Code to create the suggested figures is available at <https://github.com/nwd2918/MRS-voxel-plot>.

Voxel location

Images representing voxel locations should be produced from individual MRS voxel masks obtained from the data of all participants and transformed into standard-space. Such masks can be created automatically from most data types using a number of software tools (Edden et al., 2013; Lee et al., 2013) and are already required for partial volume correction, a necessary part of most analysis pipelines (Porges et al., 2017; Quadrelli et al., 2016). Alignment to a standard space (e.g., MNI152) can be done via individual anatomical images with any MRI analysis package, including FSL, SPM, and AFNI. The masks in standard space can then be combined to create a voxel overlap density map (Figure 1A). This shows at each standard-space anatomical voxel the percentage of overlap between the set of participant masks. With this, one can judge the consistency of voxel placement in the study and gauge the specific anatomical regions from which MRS data are being recorded. Such maps have been used in a number of previous studies (Bai et al., 2017; Duncan et al., 2019; Gaetz et al., 2014). In the case of group comparison studies separate density maps can be created and presented to help identify where there may be a systematic difference in location between the groups (Figure 1B).

Figure 1: Voxel overlap

Individual MRS voxels are converted to MNI space and then combined to show the overlap between participants. Figure (A) shows a voxel overlap density map for a single group of participants. The heatmap shows the percentage overlap of participant MRS voxels at each point in the brain. Figure (B) shows voxel overlap densities for two separate groups, as denoted by red and blue contours. The shading of the contour lines represents the percentage overlap at that point. This figure type can be extended to include multiple participant groups. Here we can see that the two groups have somewhat different patterns of voxel placement. Brain outlines are created using the Nilearn toolbox (Abraham et al., 2014).

Complementary information about voxel locations can be presented through plots of mask centres-of-mass. Here, the centre-of-mass of each participant's standard-space mask is

calculated and then displayed on a standard-space brain image (Figure 2A). Calculations can be performed in a range of tools, including, for example, FSL's `fsfstats`, python, or MATLAB. This approach is particularly useful in group comparison studies as it more clearly shows any pattern in the locations of voxels (Figure 2B). With this, any systematic differences between groups can be easily identified. Statistical analyses can also be applied to the location distributions, although this is beyond the scope of the current work.

Figure 2: Voxel centroids

MRS masks in MNI space are used to calculate the voxel centroid for each participant. Figure (A) shows this for a single group of participants. It is easy to see how tightly clustered MRS voxels are around a target anatomical location. Figure (B) shows the centroids for two separate groups, as denoted by red and blue colours. Here we can see evidence for a systematic difference in clustering between the two groups.

Figure 3: MRS spectra

Individual MRS spectra are shown along with the average spectra for a particular group. Figure (A) shows MEGA-PRESS difference spectra for a single group. Individual spectra are shown in black with the group average added in red. Figure (B) shows the same for two separate groups, as denoted by red and blue colours. Here, individual spectra are shown in the relevant group colour with the group average added as a thicker line of the same colour.

The majority of MRS analysis tools (e.g., LCModel, TARQUIN, Gannet) have options to output text files containing the processed MRS spectra. These can then be used to create figures that combine the spectra from all participants in any plotting software. Although a

group average spectrum with error bars would be an acceptable way to present this data, also showing individual spectra is optimal as this allows the reader to fully judge data quality. An example of such a plot is shown in Figure 3. This plot also demonstrates how data from multiple groups can be shown on a single figure so that data quality can be compared between them (Figure 3B). This is a key step, particularly in comparisons of patients and controls where there may be systematic differences in things such as head motion that can affect MRS data quality (Cui et al., 2018). Finally, it would be optimal to provide images of any spectra that are rejected as supplementary figures or hosted online (e.g., at <https://figshare.com/>) as this makes the spectra selection process transparent to the reader.

Conclusion

Interpretation and replication of studies requires that sufficient information be communicated to the scientific audience. An informal appraisal of the MRS literature revealed a generally low quality representation of essential information, specifically voxel locations and spectral quality. To help improve this, we set out some basic visualisation approaches that give readers sufficient information about the data included in a study to better judge its robustness and validity. Particular focus was made on studies that compare patients and controls, where there may be systematic differences in localisation and quality between groups. We hope that these suggestions and the code provided will be a useful resource that contributes to an improvement in the reporting of MRS studies.

Acknowledgements

The authors are grateful to Elizabeth McManus (University of Manchester) for useful comments on a draft of this paper. This work was supported by funding from the Taiwan Ministry of Science and Technology to NWD (105-2410-H-038-006-MY3; 108-2410-H-038-008-MY2). The authors declare no conflicts of interest.

References

- Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., Gramfort, A., Thirion, B., Varoquaux, G., 2014. Machine learning for neuroimaging with scikit-learn. *Front. Neuroinformatics* 8, 14. <https://doi.org/10.3389/fninf.2014.00014>
- Bai, X., Harris, A.D., Gong, T., Puts, N.A.J., Wang, G., Schär, M., Barker, P.B., Edden, R.A.E., 2017. Voxel Placement Precision for GABA-Edited Magnetic Resonance Spectroscopy. *Open J. Radiol.* 7, 35–44. <https://doi.org/10.4236/ojrad.2017.71004>

- Chen, J.C., Cooper, R.J., McMullen, M.E., Schriger, D.L., 2017. Graph Quality in Top Medical Journals. *Ann. Emerg. Med.* 69, 453–461.e5. <https://doi.org/10.1016/j.annemergmed.2016.08.463>
- Cui, E.D., Maddock, R.J., Buonocore, M.H., Ozturk, M., Yoon, J.H., 2018. Quantifying the effects of head movement on magnetic resonance spectroscopy estimates of gamma-aminobutyric acid. *bioRxiv* 381699. <https://doi.org/10.1101/381699>
- Duncan, N.W., Wiebking, C., Northoff, G., 2014. Associations of regional GABA and glutamate with intrinsic and extrinsic neural activity in humans: A review of multimodal imaging studies. *Neurosci. Biobehav. Rev.* 47, 36–52. <https://doi.org/10.1016/j.neubiorev.2014.07.016>
- Duncan, N.W., Zhang, J., Northoff, G., Weng, X., 2019. Investigating GABA concentrations measured with macromolecule suppressed and unsuppressed MEGA-PRESS MR spectroscopy and their relationship with BOLD responses in the occipital cortex. *J. Magn. Reson. Imaging* 50, 1285–1294. <https://doi.org/10.1002/jmri.26706>
- Gaetz, W., Bloy, L., Wang, D.J., Port, R.G., Blaskey, L., Levy, S.E., Roberts, T.P.L., 2014. GABA estimation in the brains of children on the autism spectrum: measurement precision and regional cortical variation. *NeuroImage* 86, 1–9. <https://doi.org/10.1016/j.neuroimage.2013.05.068>
- Hasler, G., van der Veen, J.W., Tumonis, T., Meyers, N., Shen, J., Drevets, W.C., 2007. Reduced prefrontal glutamate/glutamine and gamma-aminobutyric acid levels in major depression determined using proton magnetic resonance spectroscopy. *Arch. Gen. Psychiatry* 64, 193–200. <https://doi.org/10.1001/archpsyc.64.2.193>
- Hertel, J., 2018. A Picture Tells 1000 Words (but Most Results Graphs Do Not): 21 Alternatives to Simple Bar and Line Graphs. *Clin. Sports Med.* 37, 441–462. <https://doi.org/10.1016/j.csm.2018.04.001>
- Lau, J.C., Parrent, A.G., Demarco, J., Gupta, G., Kai, J., Stanley, O.W., Kuehn, T., Park, P.J., Ferko, K., Khan, A.R., Peters, T.M., 2019. A framework for evaluating correspondence between brain images using anatomical fiducials. *Hum. Brain Mapp.* 40, 4163–4179. <https://doi.org/10.1002/hbm.24693>
- Maltezos, S., Horder, J., Coghlan, S., Skirrow, C., O’Gorman, R., Lavender, T.J., Mendez, M.A., Mehta, M., Daly, E., Xenitidis, K., Paliokosta, E., Spain, D., Pitts, M., Asherson, P., Lythgoe, D.J., Barker, G.J., Murphy, D.G., 2014. Glutamate/glutamine and neuronal integrity in adults with ADHD: a proton MRS study. *Transl. Psychiatry* 4, e373. <https://doi.org/10.1038/tp.2014.11>
- Marsman, A., Mandl, R.C.W., Klomp, D.W.J., Cahn, W., Kahn, R.S., Luijten, P.R., Hulshoff Pol, H.E., 2017. Intelligence and Brain Efficiency: Investigating the Association between Working Memory Performance, Glutamate, and GABA. *Front. Psychiatry* 8, 154. <https://doi.org/10.3389/fpsy.2017.00154>
- Minshew, N.J., Goldstein, G., Dombrowski, S.M., Panchalingam, K., Pettegrew, J.W., 1993. A preliminary ³¹P MRS study of autism: evidence for undersynthesis and increased degradation of brain membranes. *Biol. Psychiatry* 33, 762–773. [https://doi.org/10.1016/0006-3223\(93\)90017-8](https://doi.org/10.1016/0006-3223(93)90017-8)
- Mullins, P.G., McGonigle, D.J., Gorman, R.L.O., Puts, N. a J., Vidyasagar, R., Evans, C.J., Symposium, C., Edden, R.A.E., O’Gorman, R.L., 2014. Current Practice in the use of MEGA-PRESS spectroscopy for the detection of GABA. *NeuroImage* 86, 43–52. <https://doi.org/10.1016/j.neuroimage.2012.12.004>
- O’Callaghan, E., Redmond, O., Ennis, R., Stack, J., Kinsella, A., Ennis, J.T., Larkin, C., Waddington, J.L., 1991. Initial investigation of the left temporoparietal region in schizophrenia by ³¹P magnetic resonance spectroscopy. *Biol. Psychiatry* 29, 1149–1152. [https://doi.org/10.1016/0006-3223\(91\)90256-I](https://doi.org/10.1016/0006-3223(91)90256-I)
- Peek, A.L., Rebbeck, T., Puts, N.A.J., Watson, J., Aguila, M.E., Leaver, A.M., 2020. Brain

- GABA and glutamate levels across pain conditions: A systematic literature review and meta-analysis of 1H-MRS studies using the MRS-Q quality assessment tool. *NeuroImage* 116532. <https://doi.org/10.1016/j.neuroimage.2020.116532>
- Porges, E.C., Woods, A.J., Lamb, D.G., Williamson, J.B., Cohen, R.A., Edden, R.A.E., Harris, A.D., 2017. Impact of tissue correction strategy on GABA-edited MRS findings. *NeuroImage* 162, 249–256. <https://doi.org/10.1016/j.neuroimage.2017.08.073>
- Quadrelli, S., Mountford, C., Ramadan, S., 2016. Hitchhiker'S Guide to Voxel Segmentation for Partial Volume Correction of in Vivo Magnetic Resonance Spectroscopy. *Magn. Reson. Insights* 9, MRI.S32903. <https://doi.org/10.4137/MRI.S32903>
- Rousselet, G.A., Pernet, C.R., 2011. Quantifying the Time Course of Visual Object Processing Using ERPs: It's Time to Up the Game. *Front. Psychol.* 2. <https://doi.org/10.3389/fpsyg.2011.00107>
- Stelzer, J., Lacosse, E., Bause, J., Scheffler, K., Lohmann, G., 2019. Brainglance: Visualizing Group Level MRI Data at One Glance. *Front. Neurosci.* 13. <https://doi.org/10.3389/fnins.2019.00972>
- Takei, Y., Fujihara, K., Tagawa, M., Hironaga, N., Near, J., Kasagi, M., Takahashi, Y., Motegi, T., Suzuki, Y., Aoyama, Y., Sakurai, N., Yamaguchi, M., Tobimatsu, S., Ujita, K., Tsushima, Y., Narita, K., Fukuda, M., 2016. The inhibition/excitation ratio related to task-induced oscillatory modulations during a working memory task: A multimodal-imaging study using MEG and MRS. *NeuroImage* 128, 302–315. <https://doi.org/10.1016/j.neuroimage.2015.12.057>
- Weissgerber, T.L., Winham, S.J., Heinzen, E.P., Milin-Lazovic, J.S., Garcia-Valencia, O., Bukumiric, Z., Savic, M.D., Garovic, V.D., Milic, N.M., 2019. Reveal, Don't Conceal: Transforming Data Visualization to Improve Transparency. *Circulation* 140, 1506–1518. <https://doi.org/10.1161/CIRCULATIONAHA.118.037777>
- Wilson, M., Andronesi, O., Barker, P.B., Bartha, R., Bizzi, A., Bolan, P.J., Brindle, K.M., Choi, I.-Y., Cudalbu, C., Dydak, U., Emir, U.E., Gonzalez, R.G., Gruber, S., Gruetter, R., Gupta, R.K., Heerschap, A., Henning, A., Hetherington, H.P., Huppi, P.S., Hurd, R.E., Kantarci, K., Kauppinen, R.A., Klomp, D.W.J., Kreis, R., Kruiskamp, M.J., Leach, M.O., Lin, A.P., Luijten, P.R., Marjańska, M., Maudsley, A.A., Meyerhoff, D.J., Mountford, C.E., Mullins, P.G., Murdoch, J.B., Nelson, S.J., Noeske, R., Öz, G., Pan, J.W., Peet, A.C., Poptani, H., Posse, S., Ratai, E.-M., Salibi, N., Scheenen, T.W.J., Smith, I.C.P., Soher, B.J., Tkáč, I., Vigneron, D.B., Howe, F.A., 2019. Methodological consensus on clinical proton MRS of the brain: Review and recommendations. *Magn. Reson. Med.* 82, 527–550. <https://doi.org/10.1002/mrm.27742>