Received: 26 July 2018 | Revised: 11 October 2018 | Accepted: 12 October 2018

DOI: 10.1002/mrm.27590

FULL PAPER

Magnetic Resonance in Medicine

# Retrospective artifact elimination in MEGA-PRESS using a correlation approach

**Sofie Tapper**[1,2] | **Anders Tisell**[1,2] | **Gunther Helms**[3] | **Peter Lundberg**[1,2]

[1]Center for Medical Image Science and Visualization, Linköping University, Linköping, Sweden

[2]Departments of Radiation Physics and Medical and Health Sciences, Linköping University, Linköping, Sweden

[3]Department of Medical Radiation Physics, Lund University, Lund, Sweden

**Correspondence**
Sofie Tapper, Linköping University, Linköping 58185, Sweden.
Email: sofie.tapper@liu.se

**Purpose**: To develop a method for retrospective artifact elimination of MRS data. This retrospective method was based on an approach that combines jackknife analyses with the correlation of spectral windows, and therefore termed "JKC."

**Methods**: Twelve healthy volunteers performed 3 separate measurement protocols using a 3T MR system. One protocol consisted of 2 cerebellar MEGA-PRESS measurements: 1 reference and 1 measurement including head movements. One-third of the artifact-influenced datasets were treated as training data for the implementation the JKC method, and the rest were used for validation.

**Results**: The implemented JKC method correctly characterized most of the validation data. Additionally, after elimination of the detected artifacts, the resulting concentrations were much closer to those computed for the reference datasets. Moreover, when the JKC method was applied to the reference data, the estimated concentrations were not affected, compared with standard averaging.

**Conclusion**: The implemented JKC method can be applied without any extra cost to MRS data, regardless of whether the dataset has been contaminated by artifacts. Furthermore, the results indicate that the JKC method could be used as a quality control of a dataset, or as an indication of whether a shift in voxel placement has occurred during the measurement.

**KEYWORDS**

artifact detection, correlation analysis, GABA, jackknife, MEGA-PRESS, MRS

## 1 | INTRODUCTION

In vivo MRS is currently used to investigate the mechanisms behind a range of neurological disorders. Determining reliable MRS concentrations is a prerequisite, and therefore essential to develop sufficiently robust techniques for data acquisition, postprocessing, and final metabolite concentration quantification of MRS data. However, MRS often requires long measurements to achieve a sufficient SNR. Therefore, MRS is sensitive to instabilities originating from factors such as subject movements, phase errors, and frequency drifts.[1] A stable data acquisition is particularly important for subtraction-based methods like MEGA-editing (Mescher-Garwood),[2-5] which is used commonly for GABA quantification, as temporal instabilities during the acquisition will be amplified by subtraction.

### 1.1 | GABA quantification

Gamma-aminobutyric acid (GABA) is the main inhibitory neurotransmitter in the mature human brain,[6] and many studies have highlighted an association between regionally altered GABA concentrations and neurological disorders.[7-12] However, there are several challenges in attempting to identify and quantify GABA, primarily due to the low SNR of the GABA signal relative to neighboring metabolite resonances. Additionally, GABA quantification becomes even more complicated because the creatine resonance signal obscures the targeted GABA signal at 3.01 ppm; consequently, robust spectral editing is needed in order to reveal a GABA resonance.

The MEGA-PRESS pulse sequence[13] is used commonly for GABA detection.[2-5,14,15] This pulse sequence generates a series of alternating spectra labeled "ON" and "OFF," in which ON spectra have been edited to alter the GABA signal. Conventionally, these ON and OFF spectra are subtracted pairwise to eliminate the overlapping creatine resonance, which generates a difference spectrum that reveals the masked GABA resonance at 3.01 ppm. Furthermore, the resulting GABA concentration is typically computed from the average difference spectrum using a quantification program, such as LCModel.[16] Finally, because no macromolecular suppression[17,18] was performed when acquiring data in this study, the term GABA+ (GABA + coedited macromolecular signal) is used throughout this paper.

### 1.2 | Artifact detection in MRS

From our experience, artifacts in the MRS data are usually expected when investigating clinical populations diagnosed with disorders such as narcolepsy, Parkinson's disease, essential tremor, or fibromyalgia. These patients often have difficulties remaining motionless during a longer acquisition. Furthermore, spectral artifacts can be observable as phase errors, frequency shifts, loss of signal amplitude, or random signal spikes in the spectrum. These artifacts influence the resulting averaged spectrum negatively, and quantification programs may misinterpret these artifacts as actual metabolite signals, which consequently result in underestimated or overestimated concentrations. Furthermore, when performing MRI studies, artifacts can often easily be detected in the final images; therefore, these images can be discarded directly. However, this practice is not equally obvious for spectroscopy data, and also due to time constraints, it is not realistic to control spectral data from each separate excitation.

Furthermore, solely for the purpose of motion correction, a prospective method (e.g., a navigator[19] or an optical system[20]) would be better for real-time tracking of subject movements. However, hardware-based implementations are often complicated and expensive. Several other techniques exist for retrospective motion correction, such as spectral registration[21] and order statistic filtering,[22,23] which have been evaluated using PRESS and MEGA-PRESS data. In our previous study using MRS datasets with motion-induced artifacts, the order statistic filtering method generated an improvement compared with the standard method of averaging the spectra; however, no significant differences were observed in the resulting concentrations computed by LCModel.[23] Finally, this result may suggest that a better retrospective method can be implemented to achieve more reliable concentrations from this MRS data.

### 1.3 | Jackknife and correlation analyses

Inspired by a previous study that used correlation of spectral windows for phase and frequency alignment of edited MRS data,[24] we implemented a technique based on an approach that combines jackknife analyses and correlation of spectral windows (Figure 1), which we termed "JKC."[25] The jackknife resampling technique is a known method for estimating the variance and bias of a data set. This jackknife estimator is calculated by systematically leaving out 1 observation from a data set, then computing 1 estimation using the remaining observations, and finally averaging over all estimations. Our hypothesis was that our retrospective approach could isolate and highlight individual artifact-contaminated MRS excitations, which would be indicated by low-resulting correlation coefficients. After removing these detected dynamic spectral acquisitions, we anticipated that the computed metabolite concentrations would be more reliable.

Finally, inspired by another clinical study at the University Hospital in Linköping, where a population diagnosed with essential tremor was investigated with cerebellar MRS, the cerebellum was chosen as the measurement region for the exploration of our hypothesis in this study. Although the cerebellum is rarely investigated using MRS, it has an important direct role in motor control as well as in "motor learning"; therefore, cerebellar MRS concentrations are of particular interest in diseases affecting motor function.

### 1.4 | Aims

The primary aim of this study was to implement an automated retrospective method for artifact detection and elimination in MRS using JKC. The secondary aim was to investigate the resulting metabolite concentrations obtained after the elimination of spectra characterized as artifact contaminated by the implemented JKC method, and to compare these concentrations with the resulting concentrations obtained when using standard averaging.

## 2 | METHODS

The Regional Ethical Review Board in Linköping, Sweden, approved this study (Dnr 2015-13-31, P. Lundberg).

### 2.1 | Data acquisitions

Twelve healthy volunteers (4 males and 8 females, mean age 29.5 years, range 22-59 years) were recruited by local advertisement, and written informed consent in accordance with the Helsinki Declaration was obtained from every volunteer. The data acquisitions have been described previously,[23] but are summarized here for convenience. All volunteers were scanned using a 3T Ingenia MR system (Philips Healthcare, Best, Netherlands) equipped with a 12-channel phased-array head coil. Before and after the spectroscopy measurements, an imaging sequence ($T_2$ weighted, turbo spin-echo sequence) was acquired to ensure correct placement of the GABA-sized measurement voxel ($3.5 \times 2.5 \times 2.5$ cm$^3$) within the left side of the cerebellum, as shown in Figure 1A. Each measurement protocol consisted of 2 MEGA-PRESS measurements (TR/TE = 2000/68 ms, editing pulse OFF/ON at 7.46/1.90 ppm, weak water suppression [MOIST; bandwidth = 140 Hz], and 320 excitations [40 dynamics: M = 20 OFF and M = 20 ON; P = 8 phase cycle steps]), in which the first spectroscopy measurement was without any intentional movements and thus considered as a reference. The second spectroscopy measurement was identical to the first, but with 4 randomized episodes of intentional head movements. Additionally, before each MEGA-PRESS measurement, a shorter unsuppressed water measurement (16 excitations [2 dynamics, 8 phase cycle steps], no water suppression) was acquired to obtain a reference for water in the tissue within the voxel. Moreover, each measurement protocol was scanned 3 times (test–retest) for each healthy volunteer, thus generating 6 distinct MEGA-PRESS measurements in which motion artifacts influenced half of these acquisitions. The motion-influenced datasets were termed dataset 1, dataset 2, and dataset 3, named according to the protocol from which the data set originated. Furthermore, dataset 1 was used as the training data for the JKC implementation, whereas datasets 2 and 3 were used for validation.

### 2.2 | Movement paradigm

Before the MRS measurements, each healthy volunteer was instructed to find the original head placement after each movement episode. The volunteer could also individually choose whether the head motion was performed in the left–right ("yaw") direction or the up–down ("pitch") direction (approximately 20º). Furthermore, the movement paradigm (example illustrated by Figure 2), thus the starting excitation and duration of each of the movement episodes, were

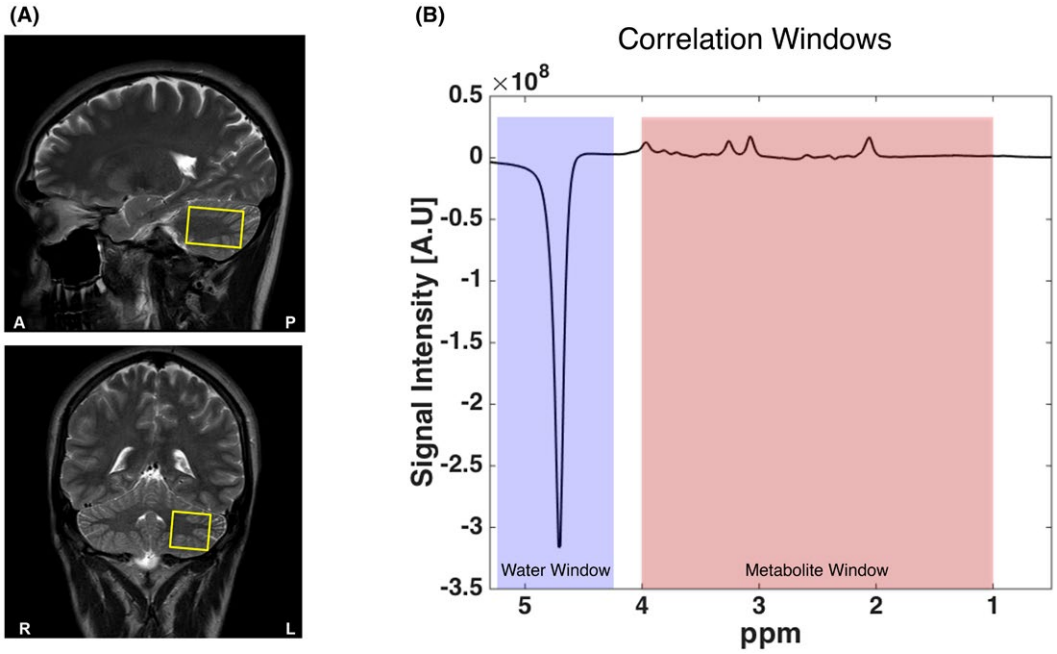


**FIGURE 1** Typical voxel placement and the correlation windows used for the method that combines jackknife analyses and correlation of spectral windows (JKC). A, The voxel ($3.5 \times 2.5 \times 2.5$ cm$^3$) was placed in the left cerebellar hemisphere. B, The metabolite window (1-4 ppm, red) and the water window (4.15-5.25 ppm, blue) evaluated in the correlation analyses. The former corresponds to the ppm range used in the metabolite concentration quantification, and the latter corresponds to the bandwidth of the water suppression. Because we used a weak water suppression method (MOIST), the water residual is clearly visible at 4.7 ppm

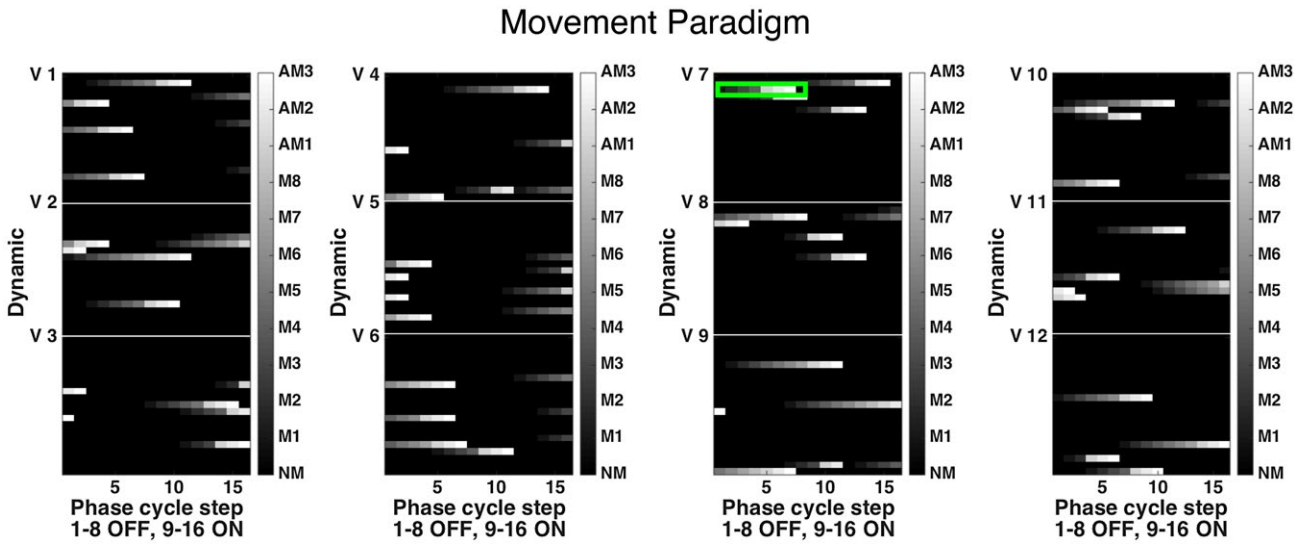


**FIGURE 2** Movement paradigm used in the acquisition of the training data. Each individual measurement consisted of M = 20 OFF and M = 20 ON interleaved dynamics, in which each dynamic consisted of 8 phase cycle steps. Each of the 12 volunteers (V) is depicted as an M × 16 matrix consisting of the M dynamics and 8 phase cycle steps for the OFF-edited and ON-edited dynamics. Thus, each "pixel" in the figure corresponds to 1 OFF/ON-edited phase cycle step, and color-coded as follows: no movement (NM) is depicted in black; movement episode (M1-M8) is depicted from dark to lighter gray; and the 3 following phase cycle steps after a movement episode (AM1-AM3) are depicted from light gray to white. The green box highlights the spectra illustrated in Figure 4

randomized using MATLAB (MathWorks, Natick MA), in which the duration could be 2-8 excitations long (4-16 seconds). The movement episodes were conducted following commands given through the communication speakers available in the scanner. After every measurement protocol, the volunteer was asked whether the movement paradigm was performed correctly. Finally, no additional hardware devices were used to monitor the actual movements during the acquisitions.

## 2.3 | Basic preprocessing of MRS data

In this project, it was necessary to access the raw data ("RAW"), which were saved separately for the dynamic spectroscopy measurements. These RAW data (lab-file and raw-file formats) consist of many more data points than the typically stored phase cycled data averages ("SPAR/SDAT" data on the Philips platform, and corresponding data formats on other scanner platforms). The RAW data correspond to the completely unprocessed measurement data, which are the output from the analog-to-digital converter. Thus, from these RAW data, it was possible to obtain the spectrum from each individual sequential excitation (i.e., phase cycle step, ON/OFF dynamic, and each separate coil element). Because the data acquisition was 512 ms in duration, the number of samples was 16 384 (sic!), as the sampling frequency was 32 kHz in the RAW data. Moreover, the RAW data were reconstructed using ReconFrame (GyroTools, Zürich, Switzerland), phase-corrected according to Klose[26] using the water reference, and frequency-aligned based on the residual water signal in the water-suppressed data.[27] After the phasing, the coil elements were combined with SNR weighting using the water reference. Finally, when using standard averaging, the phase cycle steps and dynamics were averaged to compute the resulting spectra.

## 2.4 | Jackknife correlation methodology

Figure 3 shows a workflow illustrating the whole process from data acquisition to quantification, the purpose of assisting in the explanation of JKC. First, because there was a difference in the acquisition among the 8 phase cycle steps, and between the ON-edited and OFF-edited spectra, the data were separated accordingly. Thus, data could only be compared correctly if originating from the same ON/OFF-edited phase cycle step.

The JKC method was developed using the training datasets. Two different spectral windows, metabolite and water, were used for the correlation of the spectra (Figure 1B). Consequently, only spectral points included within these 2 windows were used in the correlation analyses. The metabolite window corresponded to the analysis window (1.0-4.0 ppm), which included the targeted metabolites. The water window corresponded to the 140-Hz bandwidth of the MOIST water suppression (4.15-5.25 ppm). Additionally, because of the symmetrical placement of the ON/OFF-edited pulses around the water signal at 4.7 ppm, these ON/OFF-edited pulses equally affect the spectral points in the water window.

One jackknife correlation procedure was applied separately to each MRS measurement, each of the 8 phase cycle steps, and each type of ON/OFF editing, which in total generated M spectral windows that could be correlated accurately. Furthermore, for each of the M extracted comparable spectral windows, the spectral signal was correlated to the mean signal computed by averaging the other M-1 spectral windows. Thus, a leave-out-one analytical approach was used, which is the core of the jackknife procedure described previously. The correlation analyses were performed using the Pearson correlation, which produced a correlation coefficient. This process was repeated until all M spectral windows had been correlated to the mean of the others, which in total generated M correlation coefficients. This procedure was in turn repeated for each of the 8 phase cycle steps and ON/OFF-edited data, which in total generated an $M \times 16$ matrix for each MRS measurement that contained all resulting correlation coefficients.

## 2.5 | Finding optimal cutoff

When all correlation analyses had been performed for the training data, the next step was to find the optimal cutoff value of the correlation coefficient. The Youden index[28] was used to find this optimal cutoff, which determined under what resulting value of the correlation coefficient the spectra should be discarded, and therefore eliminated, before any further analysis. In the search for the optimal cutoff, values between 0 and 1 were tested in increments of 0.0001. The number of true positives (TPs), true negatives (TNs), false positives (FPs), and false negatives (FNs) were computed for each of the tested cutoff values, in which a TP value indicated a correctly identified movement according to the paradigm, and therefore a lower resulting correlation coefficient than the tested cutoff. Furthermore, the sensitivity and specificity were computed as TP/(TP + FN) and TN/(TN + FP), respectively. Youden's index is located

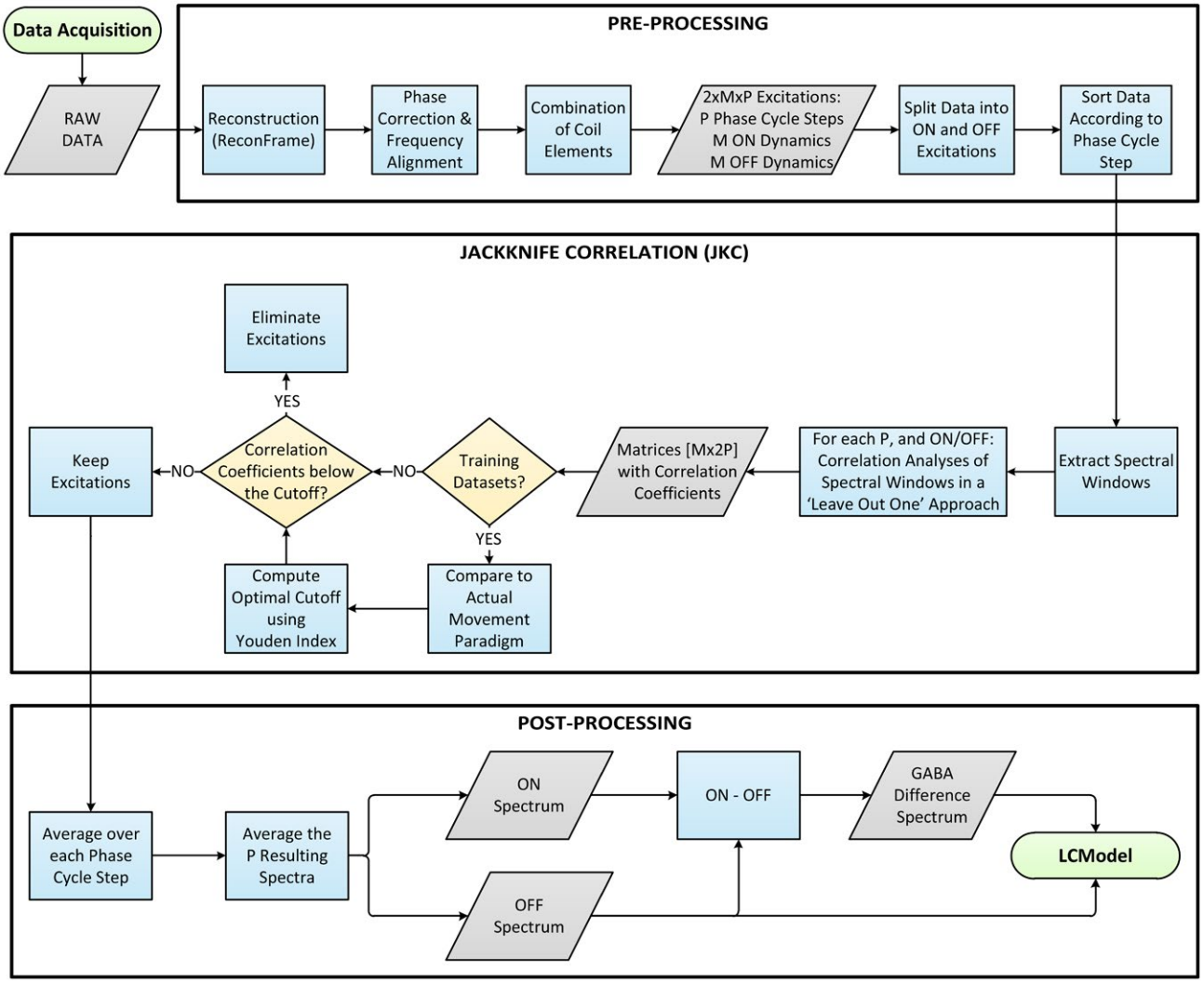


**FIGURE 3** Workflow of the different processing steps performed during the preprocessing, JKC implementation and validation, and postprocessing, before the metabolite quantification

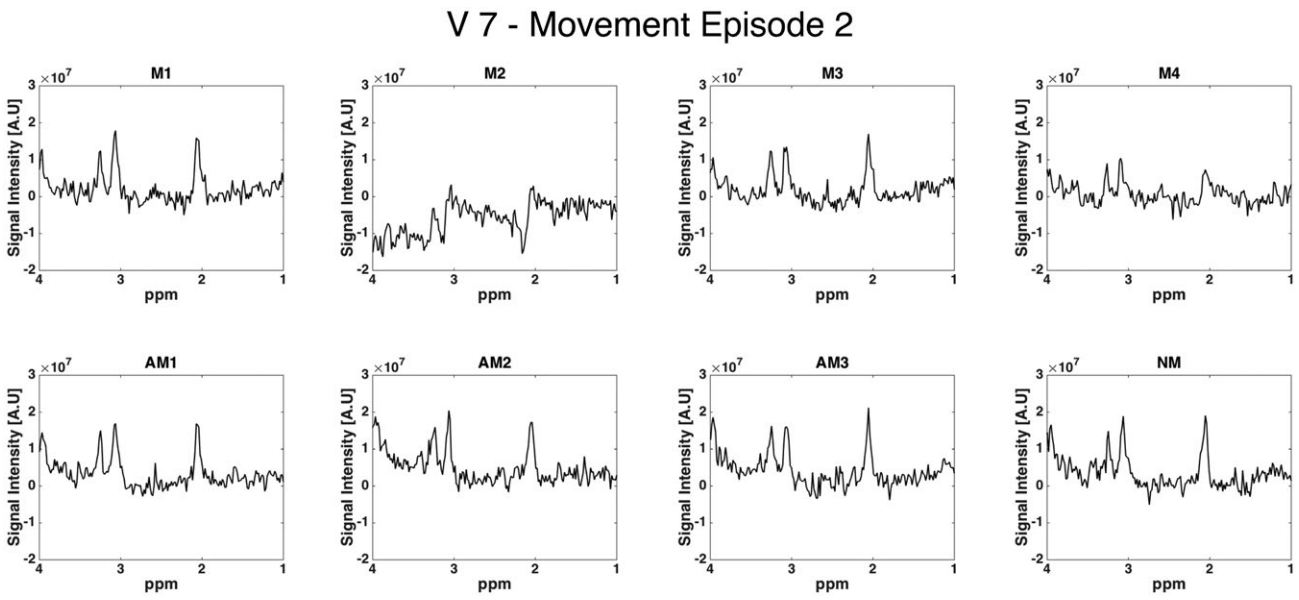


**FIGURE 4** Illustration of the effect on the spectra acquired during a typical movement episode, showing the 8 spectra indicated by the green box in Figure 2. In this example, the first spectrum in the movement episode appears to be unaffected by the movement

where (sensitivity + specificity – 1) reaches its maximum, which by definition corresponds to the optimal cutoff. Here, 4 optimal cutoffs were computed, 1 for OFF and 1 for ON, using both the metabolite and the water windows. The 4 cutoffs used in further analyses were computed using all 12 combined training data sets; however, for comparison purposes, these 4 cutoff values were also computed for each individual training data set.

## 2.6 | Quantification

The implemented JKC methodology was subsequently applied to the validation data and the reference data without any intentional head movements. Furthermore, each spectrum (acquisition) that generated a correlation coefficient below the cutoff was discarded. Then, the remaining spectra were first averaged over each phase cycle step. Finally, these 8 resulting spectra were averaged (Figure 3). The JKC method was compared with the standard technique of averaging all spectra. Furthermore, each dataset generated a resulting metabolite spectrum corresponding to the averaged OFF spectra, and a GABA difference spectrum computed by subtracting the resulting OFF spectrum from the resulting ON spectrum. The resulting spectra were then down-sampled to 16 kHz and used as input to LCModel (Version 6.3-1L).[16] The LCModel analyses were performed using the most recent basis sets obtained from the Dydak Lab,[29,30] analyzed with the main focus on the total choline (tCho), total creatine (tCr), total N-Acetyl compounds (tNA), total glutamine and glutamate (Glx), and GABA+ output concentrations, which were all referenced to water.

## 2.7 | Statistical analyses

To investigate the differences between the methodological approaches, a total of 3 paired *t* tests were performed for every computed metabolite concentration. First, the implemented JKC method was compared with standard averaging when analyzing the reference measurements. Second, the reference measurements were compared with the measurements with movement artifacts using the standard averaging method. Third, the reference measurements were compared with the measurements with movements using the implemented JKC methodology. Furthermore, the corresponding 95% confidence intervals were computed that visualize the concentration difference between the 2 groups tested. Finally, a significance level of $P < .05$ was chosen for all comparisons.

# 3 | RESULTS

## 3.1 | Movement paradigm

Figure 4 illustrates the typical effect on the spectra during a movement episode. As is evident in this figure, the first spectrum in this episode (M1) was unaffected by the head movements. Furthermore, brief delays in reacting to the vocal commands were unavoidable (Figure 5A). These delays were caused by the reaction to the scanner excitation sound by the person giving the vocal commands, and the subsequent reaction to this vocal command by the volunteer to start the movement. Thus, the movement started either during the readout phase or during the delay, which

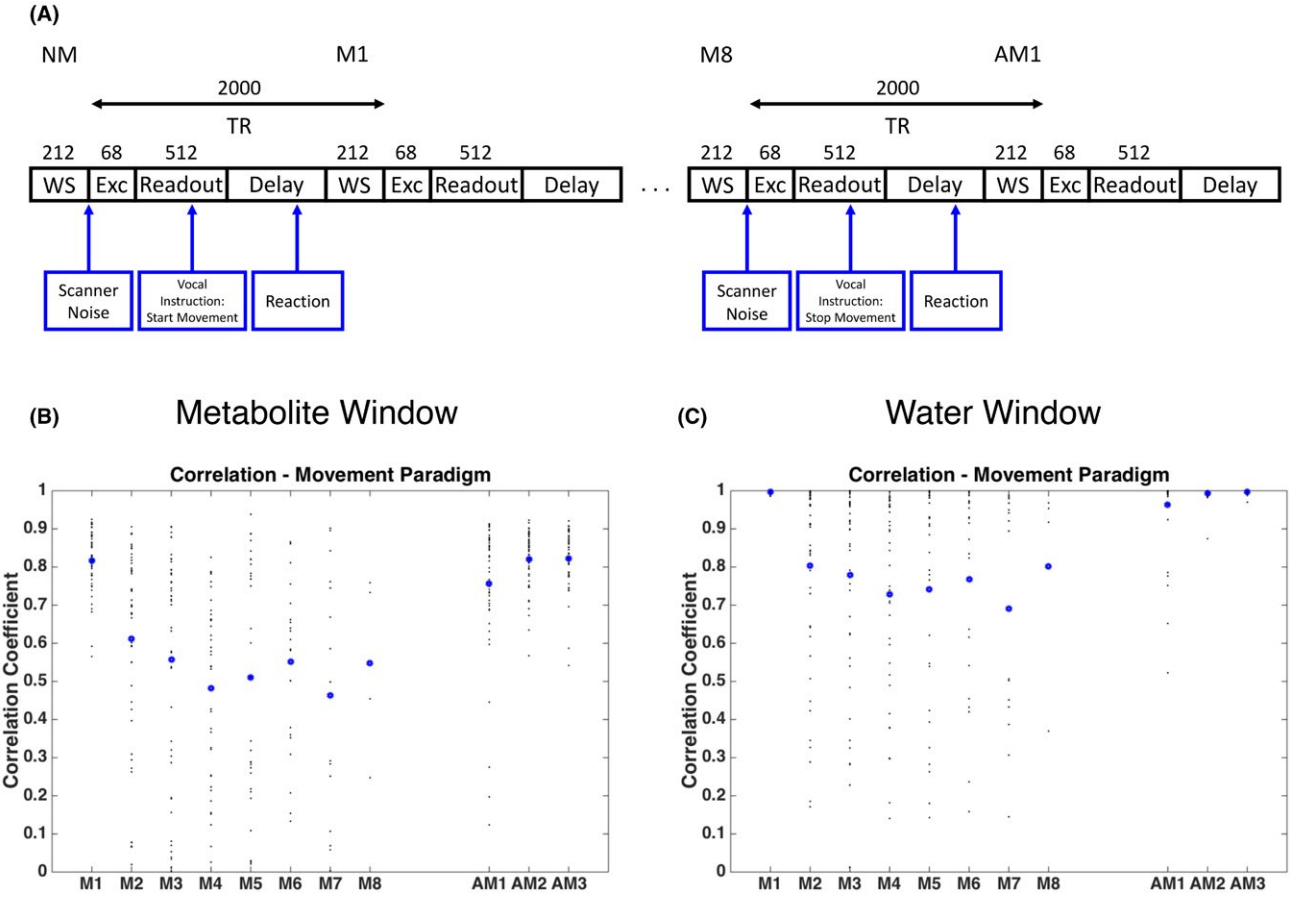


**FIGURE 5** Movement paradigm in context of the pulse sequence and correlation coefficients obtained using the 2 windows. A, The pulse sequence consists of a water suppression (WS, 212 ms), an excitation (Exc, 68 ms), a readout (512 ms), and a delay (TR = 2000 ms). The vocal instruction for the start of the movement episode was given when the excitation was performed, which was indicated by a noise from the scanner. Thus, there were 2 reaction times: 1 reaction time when giving the instruction, and 1 reaction time for the control to obtain the instruction and start the movement. The same reaction times also applied to the communication about stopping the movement. B, Correlation coefficients were computed using the metabolite window for the acquisitions containing movement (M1-M8) and the 3 following excitations after an episode (AM1-AM3). Each value corresponds to a black dot, and the mean correlation coefficient is indicated by the blue dot. C, The same as in (B) but computed using the water window

did not affect the corresponding spectrum. Additionally, this reasoning was also valid when giving the commands for stopping the head movement, and for the volunteer to find the original head placement. However, these reaction times were not long enough to also affect the excitation of the first spectrum after the movement paradigm. Moreover, the unaffected M1 spectrum was also generally observed, which was indicated by higher computed correlation coefficients for the M1 spectra (Figure 5B,C). Meanwhile, the first spectra after the movement episodes in general generated lower correlation coefficients. Both of these effects were present, regardless of which spectral window was used in the correlation analyses. Therefore, the actual movement paradigm used for the computation of the sensitivities and specificities was shifted 1 excitation in time, and therefore +2 seconds, for the JKC method to perform as well as possible.

## 3.2 | Optimal JKC method

Figure 6 shows all computed correlation coefficients for the training data using the metabolite window and the water window. The resulting correlation coefficients were much larger using the water window compared with the metabolite window. Additionally, using the metabolite window, there was a notable difference in the computed correlation coefficients between the ON-edited and OFF-edited data, which is expected due to the SNR loss in the 1.0-4.0 ppm range as a result of the ON editing.

The optimal cutoffs for the correlation coefficients were computed using the Youden index (OFF: 0.7618 versus 0.9879; ON: 0.7451 versus 0.9905), comparing the metabolite versus water windows (Supporting Information Figure S1). Additionally, the individual optimal cutoffs were computed for each of the individual training datasets (Supporting Information Table S1). The variance of the

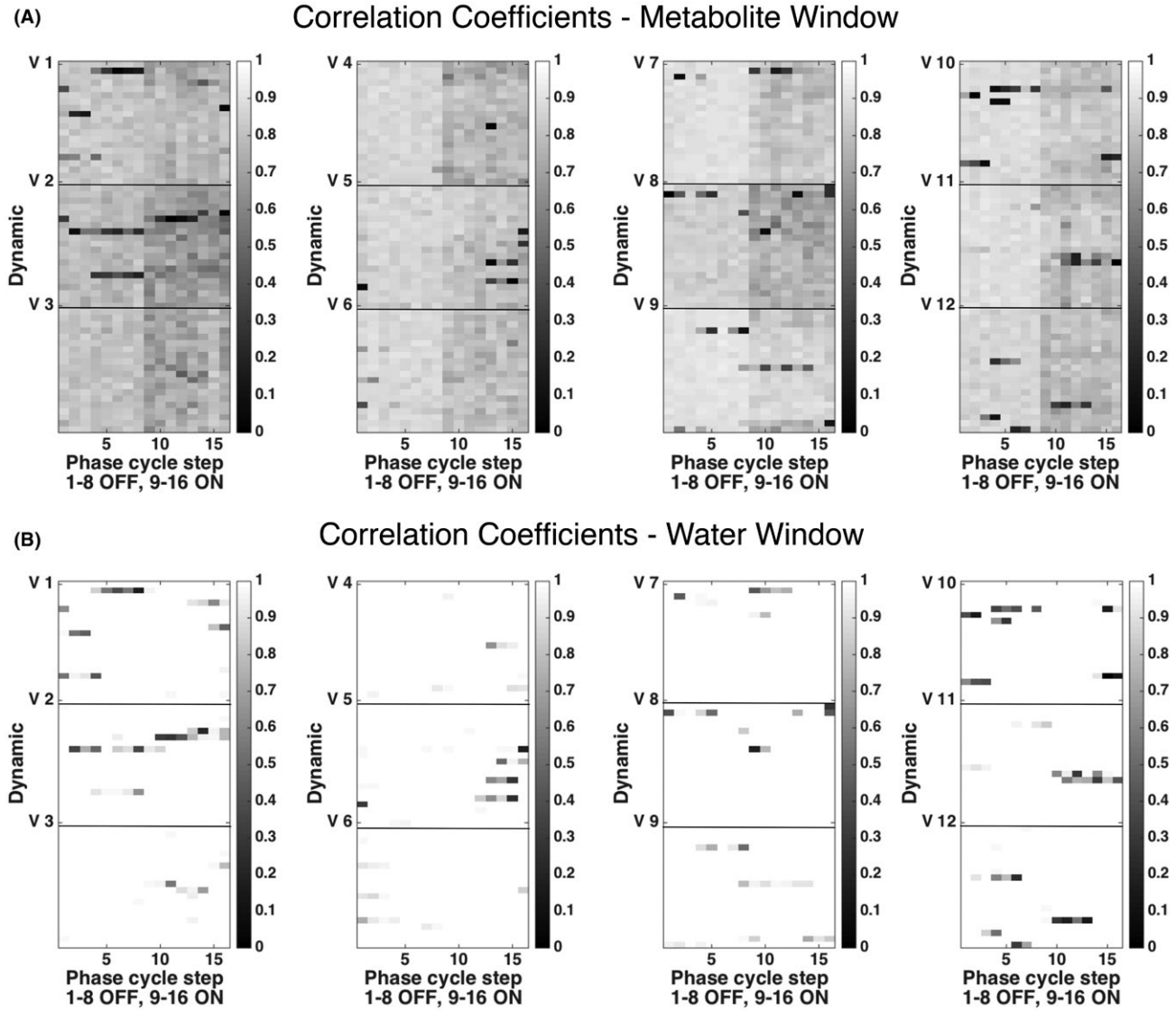


**FIGURE 6** Illustration of the computed correlation coefficients for each excitation, using the metabolite window (A) and the water window (B). Each individual measurement consisted of M = 20 OFF and M = 20 ON interleaved dynamics, in which each dynamic consisted of 8 phase cycle steps. In this figure, every measurement from each of the 12 volunteers (V) is depicted as an M × 16 matrix consisting of all excitations. Thus, each "pixel" in the figure corresponds to 1 excitation and is color-coded according to the value of the computed correlation coefficient

computed individual optimal cutoffs was much larger using the metabolite window compared with the water window. Moreover, using the water window, both higher sensitivity (OFF: 0.5455 versus 0.7190; ON: 0.7308 versus 0.7846) and higher specificity (OFF: 0.9750 versus 0.9828; ON: 0.7480 versus 0.9605) were observed than when using the metabolite window (Supporting Information Figure S2). In conclusion, the water window was more useful for detecting the movement artifacts, and was therefore used in the following analyses.

## 3.3 | Artifact detection

The artifacts detected in this study were caused by motion during the PRESS selection and the MOIST water suppression, which resulted in a combination of phase errors, frequency shifts, and magnitude errors. The developed methodology (using the water window, and the correlation coefficient cutoffs 0.9879 [OFF] and 0.9905 [ON]) was applied to the validation data, which generated the artifact detection result summarized in Figure 7. A good agreement between the movement paradigm and the detected artifacts was observed, and both high sensitivities (datasets 2: 0.7091 [OFF] and 0.7143 [ON]; datasets 3: 0.7364 [OFF] and 0.7541 [ON]) (Supporting Information Table S2) and high specificities (datasets 2: 0.9663 [OFF] and 0.9536 [ON]; datasets 3: 0.9430 [OFF] and 0.9221 [ON]) were obtained (Supporting Information Table S2). For 1 volunteer (volunteer 10) specifically, there were 2 longer patches (blue label) that were characterized as artifacts, but free from artifacts according to

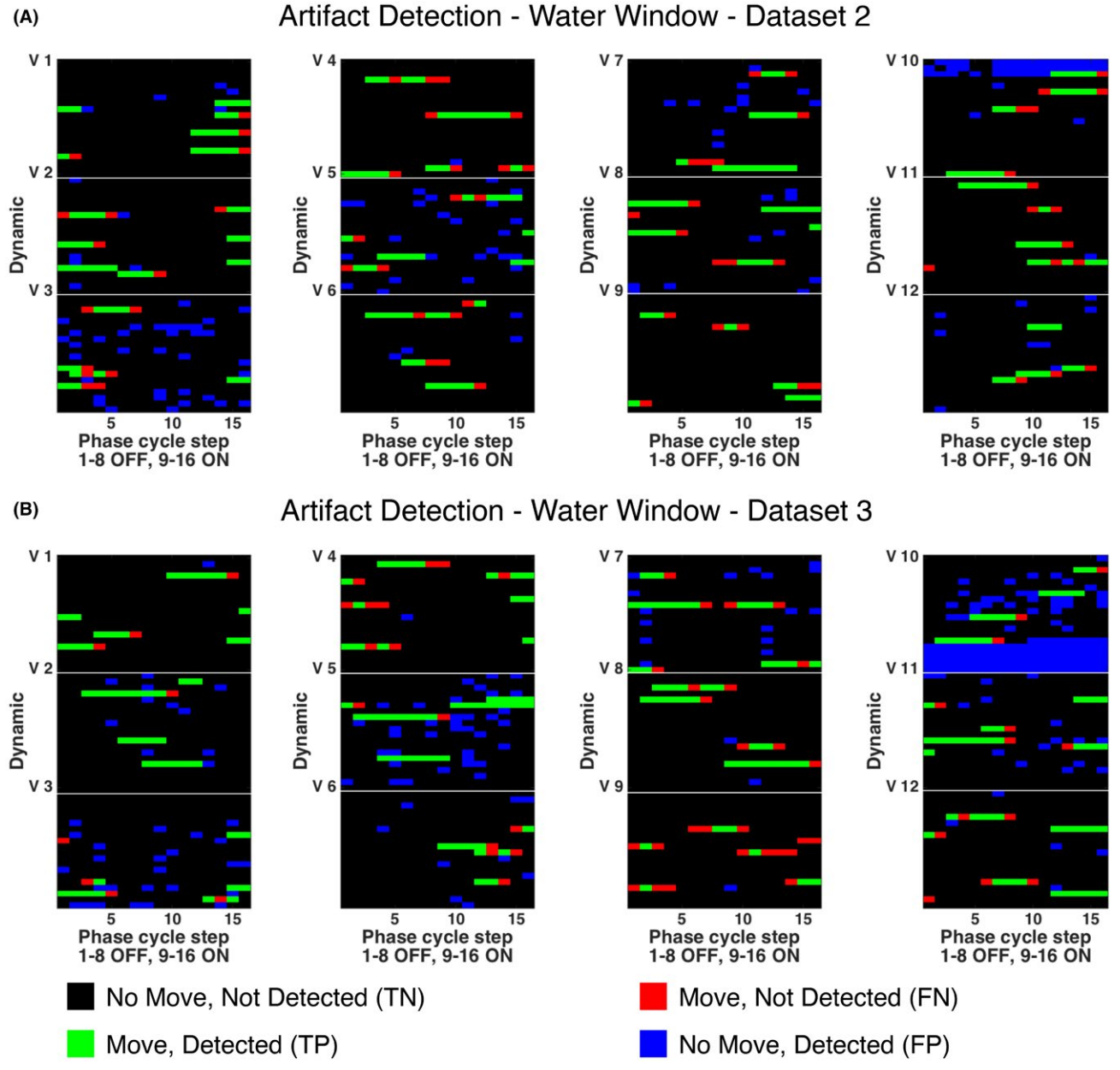


**FIGURE 7** Artifact detection using the water window applied to validation datasets 2 (A) and 3 (B). Each individual measurement consisted of M = 20 OFF and M = 20 ON interleaved dynamics, in which each dynamic consisted of 8 phase cycle steps. In this figure, every measurement from each of the 12 volunteers (V) is depicted as an M × 16 matrix corresponding to all excitations. Thus, each "pixel" in the figure corresponds to 1 excitation and was color-coded according to the result of the filtering. Black, no movement detected according to the paradigm (true negative [TN]); red, no movement detected but movement according to the paradigm (false negative [FN]); green, movement detected according to the paradigm (true positive [TP]); blue, movement detected but no movement according to the paradigm (false positive [FP])

the movement paradigm, which may indicate that a shift in the voxel position occurred.

### 3.4 | Averaged spectra and residuals

The resulting averaged spectra using JKC were computed after removal of the detected artifacts and compared with the results from standard averaging without any data elimination (Figures 8 and 9). Moreover, the 36 reference data sets without any intentional movements and the 36 data sets affected by movements were averaged separately, which generated 2 spectra that showed only small visible differences. However, looking more closely at the residuals between the mean spectra in Figures 8A and 9A, it is clear that the residuals are less apparent using the JKC methodology. Furthermore, the differences were even more obvious when the 36 individual

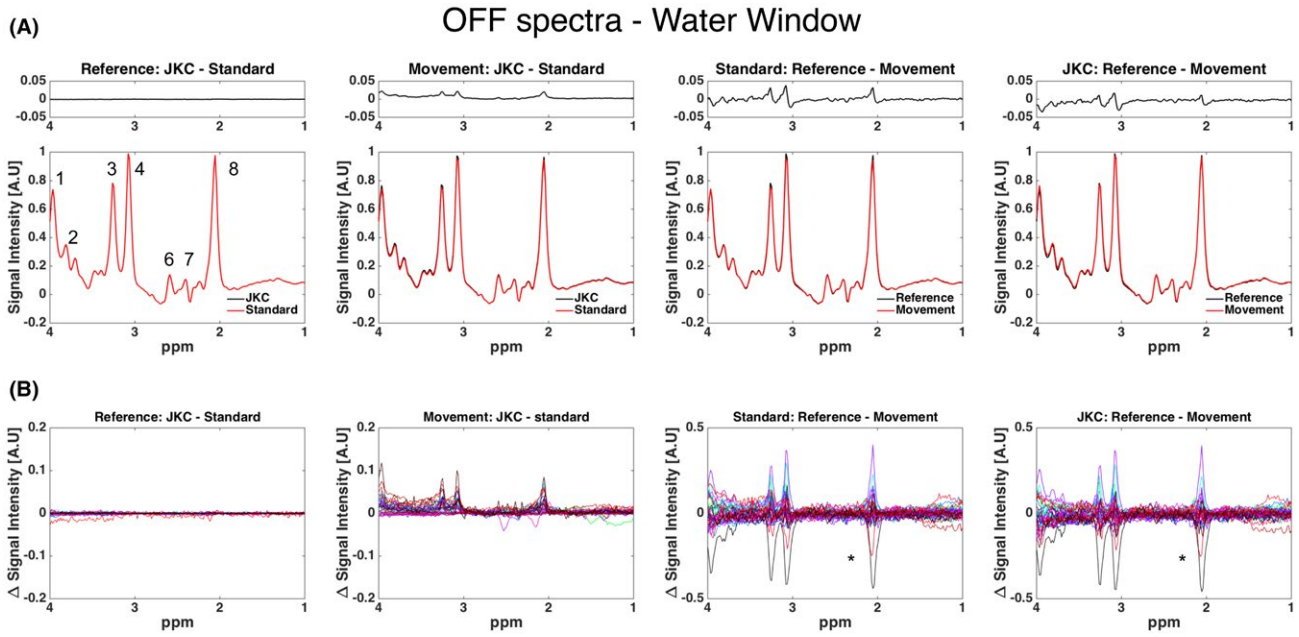


**FIGURE 8** OFF spectra and residuals obtained using the 2 different methods applied to the reference and the movement measurements. A, Two mean spectra (over 36 spectra), 1 in black and 1 in red, and the small plot above each spectrum show the residual between these 2 mean spectra. B, Residuals from the individual measurements (total of 36 in each plot). The asterisks in columns 3 and 4 in (B) highlight 2 reference measurements with less SNR than the corresponding movement measurement. 1, creatine (-2CH2-); 2, total glutamine and glutamate (Glx; -2CH-); 3, choline (-N[CH3]3); 4, creatine (-N[CH3]); 5, GABA+ (-4CH2-); 6, total N-Acetyl compounds (tNA; -3CH2-); 7, Glx (-4CH2-); 8, tNA (-2CH3)

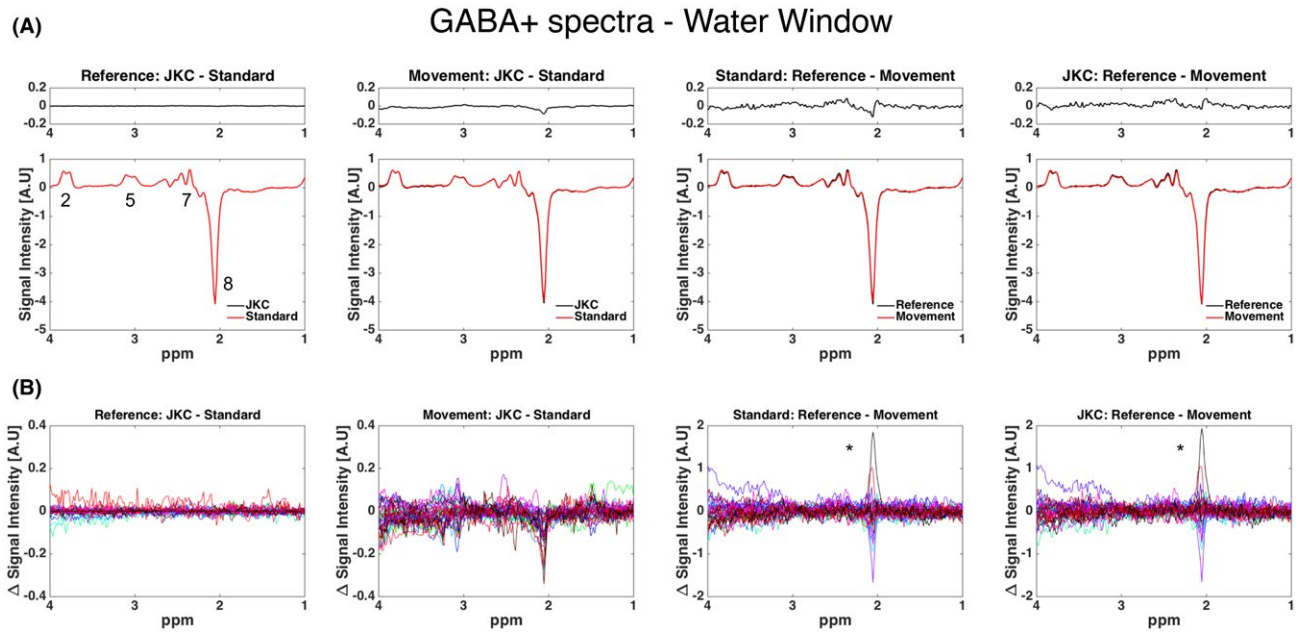


**FIGURE 9** The GABA+ spectra and residuals obtained using the 2 different methods applied to the reference and the movement measurements. A, Two mean spectra (over 36 spectra), 1 in black and 1 in red, and the small plot above each spectrum show the residual between these 2 mean spectra. B, Residuals from the individual measurements (total of 36 in each plot). The asterisks in columns 3 and 4 in (B) highlight 2 reference measurements with less SNR than the corresponding movement measurement. 1, creatine (-2CH2-); 2, Glx (-2CH-); 3, choline (-N[CH3]3); 4, creatine (-N[CH3]); 5, GABA+ (-4CH2-); 6, tNA (-3CH2-); 7, Glx (-4CH2-); 8, tNA (-2CH3)

residuals were investigated (Figures 8B and 9B). First, no large differences were observed between the JKC and standard averaging method using the reference data. In contrast, larger positive residuals were observed when investigating the OFF data that were affected by motion (Figure 8B, column 2). However, for the corresponding difference spectra (Figure 9B, column 2), this latter residual appeared to be both positive and negative in close proximity to 3 ppm, where the GABA signal was detected. Second, when comparing the movement and reference measurements obtained from the same protocol (Figures 8B and 9B, columns 3 and 4), it was apparent that the resulting residuals were both negative and positive, although more of the residuals were positive.

### 3.5 | Statistical analyses

The JKC and standard methods were compared when applied to the reference measurements, which resulted in relatively small and not significantly different confidence intervals (Figure 10, green label). Moreover, when the movement and reference measurements were compared using the standard (Figure 10, red label) and JKC method (Figure 10, black label), a larger difference was observed for the standard technique than for the JKC method. This larger difference was observed for all 5 of the targeted metabolites. Furthermore, there was a significant difference in the computed concentrations (except for Glx) between the movement measurement and the reference when the standard method of averaging was used. Those significant differences were no longer detectable for the tCho, tNA, and GABA+ concentrations when the implemented JKC method was used. Nevertheless, the *P* values were very close to the significance limit for every metabolite except tCho, which was also illustrated by the confidence intervals.

## 4 | DISCUSSION

The JKC method correctly characterized most of the spectra in the validation datasets when the actual movement paradigm was used as prior knowledge for spectral influence. After elimination of the detected artifact-affected spectra, the computed concentrations were closer to the concentrations obtained from reference datasets. Another important result was that the implemented JKC method performed equally well as standard averaging when no intentional movements were present in the data, which suggested that the JKC methodology can be used independently of whether a dataset is contaminated by motion artifacts. Importantly, the JKC method can also be used as an unbiased quality control of an MRS dataset, as a low correlation between the individual dynamic spectra is associated with a poor spectral quality. Finally, the JKC method may also be used to suggest a shift in voxel placement, which can be observed as a longer block of detected artifacts.

Furthermore, the chemical shift displacement error is a common problem when using PRESS localization at higher field strengths (3 T and above), particularly when the maximal available $B_1$ is limited.[31] One possible solution is to replace the PRESS localization with a semi-LASER (localized

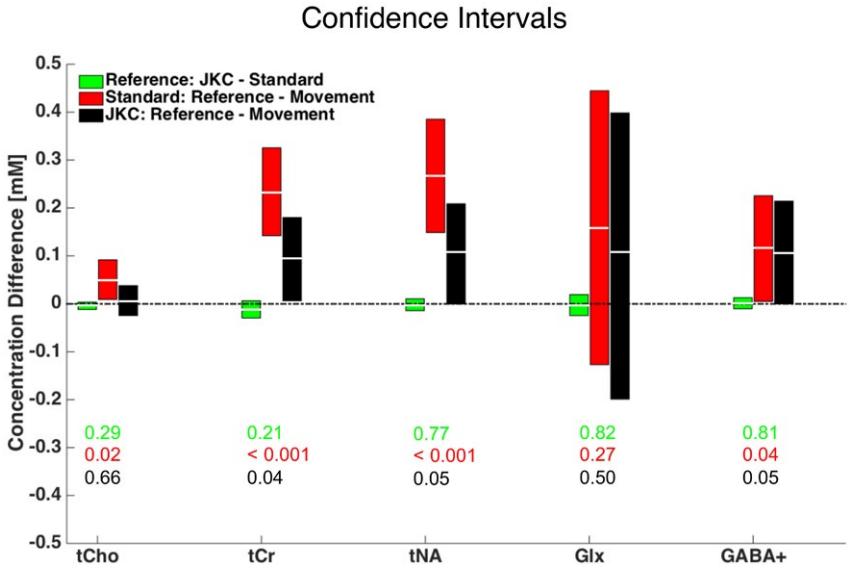


**FIGURE 10** The 95% confidence intervals obtained when performing the paired *t* tests. Green, JKC versus standard averaging applied to the reference measurement; red, reference versus movement measurements using the standard technique; black, reference versus movement measurements using the implemented JKC method. These confidence intervals were computed for the total choline (tCho), total creatine (tCr), tNA, Glx, and GABA+ concentrations. The mean concentration differences are indicated by the white lines. The corresponding *P* values are also illustrated in the figure, and color-coded according to the corresponding interval

by adiabatic selective Refocusing) pulse sequence,[32] which would minimize the chemical shift displacement error. Moreover, a change of localization technique will not affect the ability to use JKC as long as the technique involves collecting a number of averages that can be correlated.

## 4.1 | Introducing artifacts using a movement paradigm

During the correlation analyses, a 1-step time lag was observed in the spectral effect from the movement episodes compared with the intended movement paradigm. Therefore, the true movement paradigm was shifted 1 excitation step for the development of the optimal JKC methodology. This time lag cannot be explained by the expected reaction times for the volunteers to start and stop the movement, in combination with a possible temporal variation of vocal commands. We believe that this delayed effect from the movements was probably due to a $T_1$ relaxation effect, which meant that steady state was not obtained in the excitation directly after the movement stopped. Finally, instructed movement episodes are a relatively simple approach to introduce artifacts in the MRS data, but this is not the only application for the developed JKC methodology.

## 4.2 | Artifact detection using JKC

The water window was found to be superior to the metabolite window in detecting instructed motion. Thus, when using the water window, higher resulting correlation coefficients and a smaller spread in the individual optimal cutoffs were observed. This result was indeed expected due to the higher SNR of the water residual compared with metabolite resonances such as the NAA and creatine signals in the metabolite window. Using the water window, the final correlation coefficient cutoffs computed for maximal combined sensitivity and specificity were 0.9905 for ON data and 0.9879 for OFF data. The large number of decimals in the resulting cutoffs was required for optimal artifact detection, as there would have been a large stepwise difference in the results if a cutoff limit of, for example, 0.98 or 0.99 had been used. Finally, these computed cutoffs were calibrated for the present type of data, and clearly cannot be used as general cutoffs. The cutoffs may need to be recalibrated for other types of data, or even adjusted manually (i.e., like thresholding in functional MRI examination). Another interesting approach may be to eliminate a predefined number of spectra with the lowest resulting correlation coefficients prior to averaging and quantification of the MRS data.

The loss of correlation was mostly a result of a combination of deviating phases, frequency shifts, and loss of signal amplitude. This loss of correlation strength was expected, as a spatial translation in the space domain in the scanner corresponds to a phase shift in the spectral data domain. Such phase and frequency shift errors were most apparent within and around the water spectral window, due to a higher SNR of the MOIST water-suppressed residual, which readily explains the better performance of the water window. Consequently, a resulting loss of signal amplitude was apparent in the data that contained these movement artifacts (Figures 8 and 9).

Very good agreement was observed between the found artifact-influenced spectra (detected using JKC) and the movement paradigm. However, as is shown in Figure 7, there were 2 longer blue sections in the results for volunteer 10, which were incorrectly characterized as being artifacts. This suggests that the original voxel placement was not regained after a movement episode. For example, a result that could potentially be a consequence of a permanent change of the voxel position can be indicated by a different extent of macromolecular signals that contaminate the spectra, thus altering the spectral appearance. This alteration in spectral appearance will cause a loss of correlation between the spectral windows, and the JKC method will characterize spectra as artifact-contaminated, thus indicating a permanent voxel shift.

The JKC methodology labeled data as artifact-contaminated that did not contain any movements according to the paradigm. However, at the same time there were no guarantees that these data were completely free from artifacts. Nevertheless, if many spectra were characterized as artifact-contaminated, this would be a strong indication of poor general quality of the dataset. Therefore, the application of JKC may also be a suitable technique for quality control of an MRS dataset.

## 4.3 | Resulting concentrations

Compared with the standard method of averaging, we obtained a clear improvement in the computed concentrations using JKC in combination with elimination of the motion-affected spectra. The largest improvements were observed for the tCho, tCr, and tNA concentrations, which were shown by the computed confidence intervals. In contrast, the GABA and Glx concentrations were generally more uncertain, probably because of the lower SNR and their much more complex spectral patterns (multiplets instead of singlet resonances). The lack of significant improvement for the GABA+ concentrations can also be explained by both the negative and positive residuals between the spectra obtained from the 2 methods around the GABA signal at 3.0 ppm, which may suggest that this effect was a consequence of either the ON editing or the final subtraction. Finally, compared with the order statistic filtering method that we used previously for evaluating these datasets,[23] a larger improvement (smaller confidence intervals and $P$ values) was observed in the resulting concentrations when JKC was used.

### 4.4 | Limitations

We did not use any separate hardware monitoring of the movements, as this is not widely available and may be technologically challenging to use. However, there are no particular problems when using this generally applicable JKC methodology in combination with any type of navigator or device in the future. The shorter imaging sequence that was performed after each MEGA-PRESS measurement, both the reference and the movement measurement, convinced us that the MRS voxel was probably within the same planned cerebellar region during the full extent of data acquisition. Furthermore, in this project we used a GABA-sized voxel, and we do not see any major limitations using JKC when targeting smaller voxels. However, a recalibration of the cutoff is needed, which most likely would result in a lower computed optimal cutoff due to the expected SNR reduction in each spectral average when using a smaller voxel.

It was necessary to use RAW data to obtain the data from every single excitation. However, this type of data requires slightly more processing (i.e., coil combination and a subsequent combination of phase cycle steps is also necessary). The extra steps that are required in the postprocessing did not really make the processing more complex to perform, as it could be incorporated within the conventional postprocessing sequence, and it took just a few seconds longer to process. Moreover, as long as the spectra from each individual excitation can be extracted from the system, we do not see any particular reason why this methodology cannot be used on any scanner with relative ease.

Finally, the reference measurements were obviously not guaranteed to be without any artifacts. Indeed, there was an observation of 2 reference data sets that were worse in quality than most of the movement measurements (highlighted by asterisks in Figures 8 and 9). After inspection of the spectra from these 2 reference data sets, it was concluded that the volunteer had difficulties remaining motionless during the acquisition, as the similar spectral appearance when comparing the spectra to those confirmed the influence of motion artifacts. This observation strongly highlights the importance of a retrospective method for MRS quality control and artifact detection, as artifacts or poor quality obviously can influence the MRS data, not only when clinical subjects are examined, but also healthy controls.

## 5 | CONCLUSIONS

We conclude that the implemented JKC method can be applied generally to MRS data without any extra cost in the data acquisition, regardless of whether the dataset was contaminated by artifacts. Finally, our interpretation is that the JKC method can also be used as a generally applicable retrospective technique for the quality control of a dataset, or as an indication of whether a shift in voxel placement occurred during the MRS measurement.

## ACKNOWLEDGMENTS

The authors gratefully acknowledge Richard Edden for very generously providing the MEGA-PRESS pulse sequence and other tools that were developed under NIH grants P41 015909 and R01 016089. The Knut and Alice Wallenberg Foundation is also gratefully acknowledged for supporting the Seeing Organ Function project.

## ORCID

*Sofie Tapper* http://orcid.org/0000-0002-7809-2481

## REFERENCES

1. Harris AD, Glaubitz B, Near J, et al. The impact of frequency drift on GABA-edited MR spectroscopy. *Magn Reson Med*. 2014;72:941–948.
2. Mescher M, Merkle H, Kirsch J, Garwood M, Gruetter R. Simultaneous in vivo spectral editing and water suppression. *NMR Biomed*. 1998;11:266–272.
3. Mescher M, Tannus A, Johnson MO, Garwood M. Solvent suppression using selective echo dephasing. *J Magn Reson Ser A*. 1996;123:226–229.
4. Mullins PG, McGonigle DJ, O'Gorman RL, et al. Current practice in the use of MEGA-PRESS spectroscopy for the detection of GABA. *NeuroImage*. 2014;86:43–52.
5. Mikkelsen M, Barker PB, Bhattacharyya PK, et al. Big GABA: edited MR spectroscopy at 24 research sites. *NeuroImage*. 2017;159:32–45.
6. McCormick DA. GABA as an inhibitory neurotransmitter in human cerebral cortex. *J Neurophysiol*. 1989;62:1018–1027.
7. Hattingen E, Lückerath C, Pellikan S, et al. Frontal and thalamic changes of GABA concentration indicate dysfunction of thalamofrontal networks in juvenile myoclonic epilepsy. *Epilepsia*. 2014;55:1030–1037.
8. Marsman A, Mandl R, Klomp DWJ, et al. GABA and glutamate in schizophrenia: a 7 T (1)H-MRS study. *NeuroImage (Amst)*. 2014;6:398–407.
9. Chen T, Wang Y, Zhang J, et al. Abnormal concentration of GABA and glutamate in the prefrontal cortex in schizophrenia—an in vivo 1H-MRS study. *Shanghai Arch Psychiatry*. 2017;29:277–286.
10. Rowland LM, Krause BW, Wijtenburg SA, et al. Medial frontal GABA is lower in older schizophrenia: a MEGA-PRESS with macromolecule suppression study. *Mol Psychiatry*. 2016;21:198–204.
11. Emir UE, Tuite PJ, Öz G. Elevated pontine and putamenal GABA levels in mild-moderate parkinson disease detected by 7 Tesla proton MRS. *PLoS One*. 2012;7:e30918.

12. Casjens S, Dydak U, Dharmadhikari S, et al. Association of exposure to manganese and iron with striatal and thalamic GABA and other neurometabolites—neuroimaging results from the WELDOX II study. *Neurotoxicology*. 2018;64:60–67.
13. Bottomley PA. Spatial localization in NMR spectroscopy in vivo. *Ann N Y Acad Sci*. 1987;508:333–348.
14. Near J, Evans CJ, Puts N, Barker PB, Edden RAE. J-difference editing of gamma-aminobutyric acid (GABA): simulated and experimental multiplet patterns. *Magn Reson Med*. 2013;70:1183–1191.
15. Bogner W, Gruber S, Doelken M, et al. In vivo quantification of intracerebral GABA by single-voxel 1H-MRS—how reproducible are the results? *Eur J Radiol*. 2010;73:526–531.
16. Provencher SW. Estimation of metabolite concentrations from localized in vivo proton NMR spectra. *Magn Reson Med*. 1993;30:672–679.
17. Edden RAE, Puts NAJ, Barker PB. Macromolecule-suppressed GABA-edited magnetic resonance spectroscopy at 3T. *Magn Reson Med*. 2012;68:657–661.
18. Henry P-G, Dautry C, Hantraye P, Bloch G. Brain GABA editing without macromolecule contamination. *Magn Reson Med*. 2001;45:517–520.
19. Bogner W, Gagoski B, Hess AT, et al. 3D GABA imaging with real-time motion correction, shim update and reacquisition of adiabatic spiral MRSI. *NeuroImage*. 2014;103:290–302.
20. Andrews-Shigaki BC, Armstrong BSR, Zaitsev M, Ernst T. Prospective motion correction for magnetic resonance spectroscopy using single camera retro-grate reflector optical tracking. *J Magn Reson Imaging*. 2011;33:498–504.
21. Near J, Edden R, Evans CJ, Paquin R, Harris A, Jezzard P. Frequency and phase drift correction of magnetic resonance spectroscopy data by spectral registration in the time domain. *Magn Reson Med*. 2015;73:44–50.
22. Slotboom J, Nirkko A, Brekenfeld C, van Ormondt D. Reliability testing of in vivo magnetic resonance spectroscopy (MRS) signals and signal artifact reduction by order statistic filtering. *Meas Sci Technol*. 2009;20:104030.
23. Tapper S, Tisell A, Lundberg P. How does motion affect GABA-measurements? Order statistic filtering compared to conventional analysis of MEGA-PRESS MRS. *PLoS One*. 2017;12:e0177795.
24. Wiegers EC, Philips BWJ, Heerschap A, van der Graaf M. Automatic frequency and phase alignment of in vivo J-difference-edited MR spectra by frequency domain correlation. *MAGMA*. 2017;30:537–544.
25. Tapper S, Tisell A, Helms G, Lundberg P. Artifact detection using correlation analyses applied to MEGA-PRESS data containing subject head movements. In: Proceedings of the Joint Annual Meeting of ISMRM-ESMRMB, Paris, France; 2018. Abstract 2655.
26. Klose U. In vivo proton spectroscopy in presence of eddy currents. *Magn Reson Med*. 1990;14:26–30.
27. Helms G, Piringer A. Restoration of motion-related signal loss and line-shape deterioration of proton MR spectra using the residual water as intrinsic reference. *Magn Reson Med*. 2001;46:395–400.
28. Youden W. Index for rating diagnostic tests. *Cancer*. 1950;3:32–35.
29. Dydak U, Jiang Y-M, Long L-L, et al. In vivo measurement of brain GABA concentrations by magnetic resonance spectroscopy in smelters occupationally exposed to manganese. *Environ Health Perspect*. 2011;119:219–224.
30. Govindaraju V, Young K, Maudsley A. Proton NMR chemical shifts and coupling constants for brain metabolites. *NMR Biomed*. 2000;13:129–153.
31. Scheenen TW, Klomp DW, Wijnen JP, Heerschap A. Short echo time 1H-MRSI of the human brain at 3T with minimal chemical shift displacement errors using adiabatic refocusing pulses. *Magn Reson Med*. 2007;59:1–6.
32. Garwood M, DelaBarre L. The return of the frequency sweep: designing adiabatic pulses for contemporary NMR. *J Magn Reson*. 2001;153:155–177.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**FIGURE S1** Calculation of optimal cutoffs computed for OFF and ON separately, using the metabolite window (A) and the water window (B). The first column shows the receiver operating characteristic (ROC) curve with the computed area under the curve (AUC), and the maximal value of Youden's index indicated by the red line. The second column shows the Youden's index plotted against the correlation cutoff. The third column presents a zoomed version of the second column, showing the optimal correlation coefficient cutoff indicated by the maximal Youden's index

**FIGURE S2** Artifact detection for instances of dataset 1 using the metabolite window (A) and the water window (B). Each individual measurement consisted of M = 20 OFF and M = 20 ON interleaved dynamics, in which each dynamic consisted of 8 phase cycle steps. The first measurement from the 12 volunteers (V) is depicted as an M × 16 matrix corresponding to all excitations. Thus, each "pixel" in the figure corresponds to 1 excitation and was color-coded according to the result of the filtering. Black, no movement detected according to the paradigm (TN); red, no movement detected but movement according to the paradigm (FN); green, movement detected according to the paradigm (TP); blue, movement detected but no movement according to the paradigm (FP)

**TABLE S1** Computed optimal correlation coefficient cutoffs for each individual dataset 1 Note: For each dataset 1, a separate cutoff was computed for the OFF and ON data, as well as when using the metabolite window and water window. Thus, a total of 4 optimal cutoffs were computed for each data set 1, for every volunteer (V). These cutoffs can be compared with the optimal cutoffs computed using all 12 instances of dataset 1. Moreover, the spread in individual optimal cutoffs was larger when using the metabolite window compared with the water window

**TABLE S2** Summary of the computed sensitivities and specificities when using the chosen optimal cutoffs Note: The numbers indicate the total number detected from all 12 datasets (1920 excitations in total). The sensitivity was computed as TP/(TP + FN), and the specificity was computed as TN/(TN + FP)

**How to cite this article:** Tapper S, Tisell A, Helms G, Lundberg P. Retrospective artifact elimination in MEGA-PRESS using a correlation approach. *Magn Reson Med*. 2019;81:2223–2237. https://doi.org/10.1002/mrm.27590