Published in final edited form as:
Neuroimage. 2006 December ; 33(4): 1072–1081.

An Adaptive Filter for Suppression of Cardiac and Respiratory Noise in MRI Time-Series Data

Roel H.R. Deckers[1,2], **Peter van Gelderen**[1], **Mario Ries**[3], **Olivier Barret**[3], **Jeff H. Duyn**[1], **Vasiliki N. Ikonomidou**[1], **Masaki Fukunaga**[1], **Gary H. Glover**[4], and **Jacco A. de Zwart**[1]

1 *Advanced MRI section, LFMI, NINDS, National Institutes of Health, Bethesda, Maryland, USA*

2 *Biomedical NMR, Department of Biomedical Engineering, Eindhoven University of Technology, Eindhoven, the Netherlands*

3 *Imagerie Moléculaire et Fonctionnelle, ERT CNRS, Université Bordeaux 2, Bordeaux, France*

4 *Radiological Sciences Laboratory, Stanford University School of Medicine, Palo Alto, California, USA*

Abstract

The quality of MRI time-series data, which allows the study of dynamic processes, is often affected by confounding sources of signal fluctuation, including the cardiac- and respiratory cycle. An adaptive filter is described, reducing these signal fluctuations as long as they are repetitive and their timing is known. The filter, applied in image domain, does not require temporal oversampling of the artifact-related fluctuations. Performance is demonstrated for suppression of cardiac and respiratory artifacts in 10-minute brain scans on 6 normal volunteers. Experimental parameters resemble a typical fMRI experiment (17 slices; 1700 ms TR). A second dataset is acquired at a rate well above the Nyquist frequency for both cardiac and respiratory cycle (single slice; 100 ms TR), allowing identification of artifacts specific to the cardiac and respiratory cycles, aiding assessment of filtering performance. Results show significant reduction in temporal standard deviation (SD_t) in all subjects. For all 6 datasets with 1700 ms TR combined, the filtering method resulted in an average reduction in SD_t of 9.2% in 2046 voxels substantially affected by respiratory artifacts, and 12.5% for the 864 voxels containing substantial cardiac artifacts. The maximal SD_t-reduction achieved was 52.7% for respiratory and 55.3% for cardiac filtering. Performance was found to be at least equivalent to the previously published RETROICOR method. Furthermore, the interaction between the filter and fMRI activity detection was investigated using Monte Carlo simulations, demonstrating that filtering algorithms introduce a systematic error in the detected BOLD-related signal change if applied sequentially. It is demonstrated that this can be overcome by combining physiological artifact filtering and detection of BOLD-related signal changes simultaneously. Visual fMRI data from 6 volunteers were analyzed with and without the filter proposed here. Inclusion of the cardio-respiratory regressors in the design matrix yielded a 4.6 % t-score increase and 4.0 % increase in the number of significantly activated voxels.

Keywords

physiologic noise; artifacts; filtering; temporal stability; fMRI

Correspondence to: Jacco de Zwart, Advanced MRI section, LFMI, NINDS, National Institutes of Health, Bldg. 10, Rm. B1D-728, MSC 1065, 9000 Rockville Pike, Bethesda, MD 20892-1065, USA, Phone: +1 301 594 7315, Fax: +1 301, 480 1981 E-mail: Jacco.deZwart@nih.gov.

INTRODUCTION

In recent years, MRI has been successfully applied for the study of dynamic processes, such as brain activation (Ogawa 1992), cardiac function (Sakuma 1993), and tissue heating during interventional procedures (Ishihara 1995). In these studies repetitive MRI scanning is performed to generate so-called time-series data, from which the signals related to the process under study are extracted. One of the potential confounds in such studies is the contribution of quasi-periodic events that are unrelated to the process being investigated, for example the influence of the cardiac and respiratory cycles on the MRI signal amplitude during fMRI studies of brain activation. Potential artifact mechanisms are blood flow pulsations related to the cardiac cycle (CC) (Dagli 1999), and tissue motion and changes in local magnetic field related to the respiratory cycle (RC) (Windischberger 2002).

In fMRI studies, CC-related artifacts tend to be localized in areas with large vessels (Dagli 1999), whereas RC-related artifacts are often found in CSF and surrounding tissue (Windischberger 2002). Both artifacts result in a local reduction in temporal stability, compromising the detection of brain activity. This not only reduces the statistical power of fMRI experiments, but also causes temporal signal correlation of otherwise uncorrelated cortical areas, affecting interpretation of fMRI studies of event-related activity (Dale 1997) or neuronal connectivity (Biswal 1995). The need for the removal of non-white noise from fMRI data, of which CC and RC are significant sources, has been recently shown by Lund et al. (Lund 2006).

A number of methods for reducing artifacts related to CC and RC have been suggested. Provided the timing of CC and RC is known, temporal signal stability can be improved by band-rejection filtering (Biswal 1996). However, in practice, experimental design is generally such that CC and RC are temporally undersampled, complicating the separation of desired from undesired signals required to avoid suppression of the signal of interest. Also, relatively wide frequency bands would have to be discarded to account for changes in the periodicity of CC and/or RC during longer scans. Alternatively, the acquisition of navigator echoes allows correction of k-space data (Hu 1994) to reduce shot-to-shot signal variations. However, navigator correction is generally not very suitable for the suppression of RC-related, and particularly CC-related, physiological noise, due to their fine-scale spatial variation (Dagli 1999). Other work, describing an adaptive filter that derives temporal noise patterns from areas without activation, has been proposed (Buonocore 1997). This method assumes the temporal characteristics of the artifact to be similar throughout the head, an assumption which is not generally valid (e.g. due to varying latency and/or damping of local blood flow pulsation). Hu et al. (Hu 1995) proposed a method that corrects data in k-space, incorporating cardiac- and respiratory phase at the time of acquisition. A more practical alternative to this method, called RETROICOR, was subsequently suggested by Glover et al. (Glover 2000). Their method operates in image domain and models CC and RC by Fourier-series fitting.

In the following, a model-free variation of this approach is presented. Instead of model fitting, it performs selective averaging to derive an estimate of CC- and RC-related artifacts. The method is validated by evaluating its performance relative to RETROICOR.

METHODS

Filtering Method

The proposed filtering method assumes that each occurrence of the quasi-periodic disturbance (‘event’) leads to an artifact with spatial and temporal signal perturbation characteristics that only depend on the timing of the event relative to MRI data acquisition. This makes the method specifically suited for filtering of MRI data obtained with a single-shot acquisition. The method

is not suitable for filtering data whose acquisition time per image approaches or exceeds the time scale of the event (e.g. most multi-shot data). The timing of the event is typically asynchronous with the MRI acquisition, causing a deterioration of the MRI time series data. If a sufficiently long time-series is acquired, a precise estimate of the temporal characteristics of the artifact can be obtained with a temporal precision that exceeds the temporal resolution of the MRI data by a large factor. This is achieved by sorting and averaging of MRI data based on their acquisition time relative to the nearest 'event' (Figure 1). The period encompassing each event is subdivided in a number of discrete intervals ('bins'). All MRI data with acquisition times that fall within the same bin (meaning that they are in the same interval of acquisition time relative to the event) are averaged on a voxel-by-voxel basis. This selective averaging results in a suppression of signals whose timing does not correlate with the event (including random noise), while signals that correlate with the event are preserved. The resulting data are referred to as the 'artifact estimate', which are four-dimensional (3 spatial dimensions and one temporal dimension). Subsequently, the baseline signal is removed from the artifact estimate on a voxel-by-voxel basis. The baseline signal is defined as the temporal average of the artifact estimate. Removal of the artifact from the data should therefore not lead to a net change in the signal intensity over time. As a final step, the artifact estimate is subtracted from the MRI time series data. This is done by subtracting the bin in the artifact estimate from all MRI data that were initially assigned to this bin based on their acquisition time relative to the event.

An important parameter that affects performance of the method is the number of bins used for averaging. The use of a large number of bins leads to an artifact estimate with greater temporal accuracy. On the other hand, precision is reduced due to reduced averaging, and there is a larger reduction of the number of degrees of freedom of the experiment. The optimal number of bins therefore is a compromise between accuracy and precision and will depend on the experimental conditions.

Here, the length (duration) of the artifact estimate (and corresponding bin duration) is determined from the RC and CC timing data. First, the mean and the standard deviation of the interval between events are computed for CC and RC separately. The maximum interval, while excluding intervals that exceed the mean plus two standard deviations, is used as the overall duration of the artifact estimate. Thereafter, the bin duration follows from division of the overall artifact estimate length by the number of bins.

Discarding the outliers based on standard deviation avoids the use of an excessively large artifact estimate length in case of a short breath hold by the volunteer or a missed event (e.g. cardiac trigger pulse). This will potentially lead to some MRI acquisition time points being outside of the filtering range. These time points will not be corrected. In addition, when a given bin contains less than four volumes, the limited number of averages in this bin would potentially cause the introduction of additional noise in the MRI time series data. These volumes are therefore also not corrected. The number of uncorrected volumes due to these two effects is however very limited. In the experiments described here the percentage of uncorrected volumes did never exceed 2.5 % of all available time points.

Experimental Setup

Performance of the filter was evaluated using data acquired on a General Electric 3 T scanner (GE Healthcare, Waukesha, WI, USA) on 6 normal volunteers (3 female, 3 male, aged 24–52 years, 31.8 years old on average). The volunteers gave informed consent to a research protocol approved by the institutional review board (protocol number 00-N-0082, latest amendment approved by the IRB on March 15, 2005). Two gradient-echo EPI scans were performed on each of the volunteers, one with a TR of 1700 ms, which resembles a typical fMRI experiment and is referred to as 'slow', and a second single-slice dataset with a TR of 100 ms, referred to

as 'fast'. The 'slow' dataset was acquired with the following parameters: 17 slices; 90º flip angle; 366 volumes (622.2 s total acquisition time); 45 ms TE; 2.3×2.3×3.0 mm^3 nominal spatial resolution (96×72 acquisition matrix and 3 mm slice thickness). The single slice of the 'fast' dataset was spatially aligned with the center slice of the 'slow' dataset. The slice (and consequently the slice stack for 'slow') was oblique-axial, approximately parallel to and at the level of the calcarine fissure. Scan parameters for 'fast' were identical to 'slow', except for the following: single slice; 100 ms TR; 15º flip angle; 6222 acquired volumes. This 'fast' scan was performed to ensure the unaliased acquisition of both CC- and RC-related artifacts, which facilitated proof of principle and performance evaluation. To assure steady state of the signal, the data acquired during the first 8.5 seconds of scanning (5 volumes of 'slow' and 85 of the 'fast' dataset, respectively) were discarded. To boost SNR, data were acquired using a 16-channel head coil (de Zwart 2004), which was connected to a custom-built 16-channel receiver (Bodurka 2004). Data from the 16 coil elements were combined as described previously (de Zwart 2002). Note that sensitivity-encoded parallel imaging was not used, so an acceleration rate (SENSE-factor (Pruessmann 1999)) of 1 was used for reconstruction. The result is similar to root-sum-of-squares combination of the coil signals, but the reconstruction is phase-sensitive and additionally accounts for noise correlation between coil elements. After reconstruction, magnitude images were registered to the last image in the time series to reduce possible effects of bulk motion using C-based software developed by Thévenaz et al. (Thévenaz 1995).

Performance of the new method was evaluated by calculating the reduction in temporal standard deviation of the signal as a result of filtering. To reduce potential contribution of slow drifts (e.g. drift of the magnetic field), 8^{th}-order polynomial detrending was performed prior to filtering. Note that, in case of the 'slow' data, this might potentially eliminate a small fraction of the aliased CC- or RC-related signals, depending on the frequency of the aliased artifact.

CC and RC were recorded using a pulse oximeter (placed on the index finger) and respiratory bellows, respectively, both of which are part of the General Electric MRI scanner. The scanner provides an interface through which transistor-transistor logic (TTL) pulses for CC can be obtained. These pulses coincide with the top of the peak that results from the cardiac QRS-wave complex. Secondly, the interface provides an analog output of the respiratory bellow. Finally, the scanner can provide user-programmable TTL pulses through this interface, which in our case were given at the time of each excitation pulse (except for the first volunteer, where such pulses were given only every 8^{th} TR and subsequently interpolated to obtain a timing reference for each shot). These three signals were digitized using a National Instruments DAQCard-AI-16E-4 (National Instruments Corp., Austin, TX, USA) and software that was developed in-house, running on a PC using the Windows XP Professional SP2 operating system. A sampling frequency of 1 kHz was used for each of the three channels. The acquisition was started by the first MRI trigger pulse and ran for 640 s (18 s longer than the MRI acquisition time). MRI trigger pulses were recorded to allow correction for slight timing differences between the acquired physiologic signals and MR data, since the clocks of the MRI and acquisition PC were not synchronized. The top of each respiratory wave was chosen as the marker for each respiratory cycle.

Filtering Performance Evaluation

The filtering method was implemented in IDL (Research Systems, Inc., Boulder, CO, USA), and processing was done on a dual AMD Athlon MP based workstation running SuSE Linux 9.3. To determine the optimal number of bins to use for filtering, the 6 (detrended) 'slow' datasets were filtered using values for the number of bins ranging from 4 to 99. The 'slow' datasets were used, since these are more representative of the kind of scan commonly performed in the laboratory than the 'fast' scans. To increase sensitivity, only voxels that had a significant artifact contribution were taken into account. The amplitude of the artifact peak in the

magnitude frequency spectrum of the corresponding 'fast' dataset was used as a measure of the severity of the artifact for each pixel. The noise estimate was based on the standard deviation of the last 256 samples in the spectrum (the frequency range from 4.58 to 5.00 Hz). Subsequently, the average amplitude in the magnitude of the spectrum was computed over a frequency range of 0.15 Hz surrounding the principal artifact frequency as determined from the timing data. This range was chosen to match the average full-width-at-half-maximum (FWHM) of the cardiac artifact (the FWHM for the respiratory artifact was 0.10 Hz on average over the 6 volunteers). Only pixels in which a substantial artifact was present were included in the performance analysis. Voxels were included if the average intensity in the respective range (either cardiac or respiratory) in the magnitude spectrum exceeded 1.7 times the spectral baseline SD. Two separate runs were performed, one in which data were only filtered for artifacts related to CC, the other only for RC-related artifacts. As a measure of artifact suppression, the improvement in temporal standard deviation (SD_t) compared to unfiltered data was computed for each of the resulting datasets. Relative SD_t was defined as SD_t after filtering divided by SD_t before filtering, corrected for the decrease in degrees of freedom incurred by filtering. Only data from the center slice of 'slow' for all six volunteers were used, since this slice corresponds to the 'fast' data that were used for voxel selection. For all 6 volunteers combined, the selections comprised of 864 voxels for cardiac and 2046 voxels for respiratory filtering. Of these voxels, 271 were included in both the cardiac and respiratory selection. Note that filtering was performed on all voxels. These subsets were selected only to evaluate the performance of the method.

In order to further demonstrate performance of the method for the filtering of undersampled data, the 'fast' dataset was split into 17 subsets, all of which were filtered independently. Each subset contained 361 volumes (the first 85 volumes of 'fast' were discarded, see above), which were acquired 1.7 s apart (equal to 'slow'). The first subset contains the volumes [1, 18, 35, …, 6121], the second subset the volumes [2, 19, 36, …, 6122], etc. After filtering, the data were combined in the order they were originally acquired in so that the performance could be evaluated in the same way as the 'fast' data filtered as a complete set.

Comparison to RETROICOR

C-code with the RETROICOR implementation described in (12) was obtained from the Center for Advanced MR Technology at Stanford (http://rsl.stanford.edu/research/camrt.html). Since the implementation was designed for the analysis of short-integer data and a square image matrix, reconstructed magnitude data were zero-filled to 96×96 matrix size and intensity was subsequently scaled so that the maximum intensity of the data was about 10% below the maximum value for signed 16-bit integers (32767). This was done to minimize the amount of noise that is introduced by rounding errors while avoiding the risk of clipping. The RETROICOR implementation was not designed to handle timing data acquired with an external digitizer with its own clock, so there was no input for MR trigger timing. To account for clock jitter between MRI data and the data from the physiologic monitoring computer, the signal from the respiratory bellow for each interval between subsequent MRI pulses was resampled based on the MR trigger pulses acquired with the physiological noise computer. The respiratory data were subsequently downsampled to 40 Hz for this RETROICOR implementation. The time of each cardiac trigger pulse was compared to the time of the nearest MR trigger pulse and then corrected for the difference between the detected MR trigger time and the MRI acquisition time derived from the scanner. After RETROICOR-filtering, the data were scaled back to their original signal level and processed identically to the data filtered with the method described here.

Effect of Filtering on fMRI

Simulations were performed to investigate the effect of filtering on fMRI data analysis. A simulated fMRI time course was generated with 366 time points and a 1.7 s TR (similar to 'slow'). The signal-to-noise ratio was set to 100 by adding computer-generated normally distributed noise. To simulate activation, a 30 s on, 30 s off, block stimulus was convoluted with a truncated Gaussian hemodynamic response function with 3.5 s width and 3.5 s latency (Waldvogel 2000), and added to the simulated baseline signal. The activation signal amplitude was set to 1 % of the baseline signal level. Monte-Carlo simulations were performed by filtering this time course for cardiac artifacts using heart rates ranging from 40-120 beats-per-minute (BPM) and a number of bins ranging from 4–80, both varied in increments of 1. Noise was added to the heart rate in the form of random jitter derived from randomly distributed noise with a standard deviation of 1 % of the heart beat interval. For each set of parameters, 256 time courses were generated for averaging purposes.

Note that no actual simulated cardiac artifact was added to the data, but that only the timing of this artifact was simulated. This allowed assessment of the effect of physiological artifact filtering on the measured fMRI signal amplitude.

The simulated data were analyzed either by applying the filter before fMRI regression analysis (sequentially), or by combining the filter and hemodynamic response function into a single design matrix, thereby filtering and analyzing simultaneously. Simulations with the sequential method were performed twice, once with our method and once with RETROICOR (in cardiac-only mode) for the filtering step. The simultaneous approach was performed with a variation of our method only. Each bin of the cardiac filter was described by a separate regressor: a function that was '1' for each time point that was an element of that specific bin, and a '0' for all other time points. These regressors where then added to the design matrix, which also contained the hemodynamic response function for fMRI analysis. The resulting regression analysis yielded both fMRI signal change amplitude information as well as a measure of the artifact level in each bin. As a result, cardiac artifacts were removed from the residual signal, which would subsequently be used to compute the t-score.

Functional MRI data from an unrelated study were used to demonstrate the benefit of filtering performance in an actual fMRI experiment. These data (n=6) were acquired on six normal volunteers under the same IRB-approved protocol, with the same hardware and using similar settings as the 'slow' scan. Experimental differences compared to the 'slow' scan were the use of a slightly smaller FOV (210×158 mm^2) and a slice thickness of 2 mm, resulting in a nominal voxel size of 2.2×2.2×2.0 mm^3. Furthermore, only 10 slices were acquired with a 1 s TR and a 70° flip angle. A 5-minute, 30-s off/30-s on block paradigm was used (5 blocks). During on-periods a 7.5 Hz contrast-reversing radial checkerboard stimulus was shown. The stimulus covered approximately the central 21 degrees of the visual field. Image registration and detrending were performed as described earlier. Six hundred volumes were acquired during the 10-minute scan, but only volumes 31–330 were analyzed for this experiment. Analysis was performed twice, once using only a regressor for baseline signal intensity and one for the block paradigm stimulus convoluted with the with a truncated Gaussian hemodynamic response function with 3.5 s width and 3.5 s latency (Waldvogel 2000) ('unfiltered' design matrix). Secondly, analysis was performed using an expanded design matrix which contained the two 'unfiltered' regressors and to which 40 cardiac regressors and 21 respiratory regressors were added as was described above ('filtered' design matrix). The IDL code used for analysis (Waldvogel 2000) detected (Durbin 1971) and corrected (Watson 1955) serial correlation in the regression analysis. Voxels in which the t-score was equal to or larger than 5 in either of the two analyses were included in a selection of activated voxels.

RESULTS AND DISCUSSION

Determining the Optimal Number of Bins

A plot of the reduction in temporal standard deviation as a function of the number of bins used for filtering is shown in Figure 2. Data shown represent the average over the voxels that were identified as suffering from RC- or CC-related artifacts (based on spectral analysis, see *Filtering Performance Evaluation* in METHODS), respectively 2046 and 864 voxels. The data shown in red demonstrate the average SD_t reduction after cardiac filtering as a function of the number of bins, data in blue are the result for respiratory filtering. The solid lines in Figure 2 show the result of boxcar averaging with a window size of 16 points, used to suppress high-frequency fluctuations. The highest SD_t-reduction (lowest relative SD_t) found by determining the minimum of these smoothed data, respectively 40 bins for cardiac and 21 for respiratory, was used as the number of bins used in further analysis.

The optimal number of bins depends on a number of parameters. On one hand, bins need to be sufficiently short, and thus the number of bins substantially large, to be able to adequately represent the fluctuations in the MRI data that are the result of the events. On the other hand, bins need to be sufficiently populated to average out signals not related to the event. A smaller number of bins is therefore expected to give better filtering performance when the event does not introduce fluctuations at a rate that is high relative to the artifact interval. This can explain why the optimum number of bins is lower for RC than for CC. The volunteers were at rest in the scanner, and their breathing was relaxed, causing predominantly low-frequency fluctuations in the MRI signal that can be covered by a smaller number of longer bins. Since CC is a fast contraction of the heart muscle followed by a short period of rest, shorter bins are required to adequately represent the resulting perturbation of the MRI signal.

Level of Artifact Suppression

The detected respiratory and cardiac rates for the different volunteers are shown in Table 1. Comparison of the rates found for the ‘slow’ and ‘fast’ dataset shows that both cardiac and respiratory rate were stable from run to run, but varied substantially from volunteer to volunteer. The bin duration for filtering was derived from these physiological data and is also shown in Table 1.

An example of the filtering performance for CC- and RC-related artifacts is demonstrated in Figure 3 for a pixel in a ‘fast’ dataset. It shows the magnitude frequency spectrum for a voxel that had both a substantial cardiac and respiratory artifact. As can be seen in the top-left plot (labeled ‘unfiltered’), CC resulted in an artifact peak at approximately 0.8 Hz (with distinguishable harmonics at 1.6, 2.4 and 3.2 Hz). This is supported by the CC timing data obtained from the pulse oximeter, which detected an average heart rate of 48 beats-per-minute (BPM) (0.79 Hz) for this volunteer. A respiratory artifact peak can be distinguished at approximately 0.25 Hz. During the acquisition of the ‘fast’ data, measured average heart rates ranged from 0.79-1.33 Hz (average: 1.01 Hz) for the 6 volunteers. Average respiratory frequencies for the volunteers ranged from 0.25-0.36 Hz, with a mean of 0.30 Hz. The plot in Figure 3 labeled ‘cardiac filtered’ shows the spectrum after filtering for CC only, demonstrating a substantial reduction of the CC artifact peak, as well as its harmonics, without quantifiable effect on the remainder of the spectrum (e.g. the respiratory peak at 0.25 Hz). The plot labeled ‘respiratory filtered’ shows similar data for RC-only filtered data, indicated by suppression of the peak at 0.25 Hz. Respiratory filtering applied to the cardiac-filtered data (shown in the plot labeled ‘cardiac and respiratory filtered’) shows a cumulative effect. Changing the order, applying cardiac filtering to respiratory-filtered data, leads to similar results (data not shown). Note that the noise level in the frequency ranges from which peaks were removed by filtering is similar to the noise level in surrounding areas, demonstrating that this filter does not act as

a band-reject filter. Filtering the data in 17 downsampled subsets (see METHODS), in which CC and RC were undersampled, resulted in similar performance, as is shown by the blue line in Figure 3. This demonstrates that the method is as effective for undersampled artifact signals as it is for a sufficiently sampled artifact.

The number of bins used to filter the data shown in Figure 3 was derived from the optimization for the 'slow' dataset. The 'fast' data shown in Figure 3 contain 17 times more time points (albeit with a lower SNR, due to lower flip angle and shorter TR, and presumably higher temporal correlation), so they could potentially benefit from filtering with a different number of bins. Since the 'fast' experiment was used to demonstrate feasibility of the method only, the optimal number of bins for this scan was not investigated.

The spatial distribution of the CC- and RC-related artifact, as well as the level of artifact suppression, is shown in Figure 4. Maps of CC- and RC-related artifact level both before and after filtering are shown for the 'fast' dataset acquired on one of the volunteers. Artifact level is computed as was discussed above, by averaging spectral intensity in the magnitude frequency spectrum over a range of 0.15 Hz surrounding the primary artifact frequency (for this volunteer respectively 0.28-0.43 Hz for respiratory and 1.26-1.41 Hz for cardiac). The average intensity is shown in units of noise level (defined as the standard deviation derived from the frequency range 4.58-5.00 Hz). The right-most column shows the artifact map after filtering the data in 17 subsets, showing similar (albeit somewhat inferior) filtering performance as the data filtered as a whole (center column).

Filtering performance was similar for all six volunteers, as is shown in Figures 5 and 6 for all voxels with respectively CC- and RC-related artifacts. The reduction of the average spectral intensity surrounding the principal cardiac artifact peak resulting from filtering using the method proposed here is shown in Figure 5A for all 864 voxels that had a substantial cardiac artifact. Data are sorted based on pre-filtering artifact intensity (dashed line) and shown as the average intensity in units of noise level standard deviation (derived from the spectral range 4.58-5.00 Hz). The artifact intensity after filtering is shown as a solid line. Ideally, when the artifact is completely removed, the expected value for the average spectral intensity is 1, assuming that no other signals are present in that frequency band. Similar results for the 2046 voxels with a substantial respiratory artifact are shown in Figure 6A. These data demonstrate that the artifact is approximately suppressed to noise level in the majority of the cardiac voxels and a substantial number of the respiratory voxels. In several cases filtering leads to only minor improvements. This is more commonly the case for respiratory filtering than cardiac filtering (see Figures 5 and 6). One possible contributor to this effect is the more limited number of respiratory events than cardiac events (187 versus 646 events on average during the 640 sec physiologic data were sampled). In addition, respiratory artifacts are present in a lower range of spectral frequencies than cardiac artifacts. This spectral region that is more likely to be affected by scanner drift and functionally-related low frequency signal fluctuations (Biswal 1995), signals that are unaffected by the filter since they are asynchronous with respiration. The presence of such low-frequency signals, as well as the fact that they are not removed by the filter, can be seen in Figure 3. Presence of such signals would mimic reduced filtering performance the way this performance is expressed in Figure 5 and 6 since the residual signal level in that part of the spectrum would deviate substantially from noise level. The similarity of these and the RETROICOR results (see below) supports this proposition.

Finally, the average reduction in SD_t was computed for the selected voxels, both in the 'fast' datasets and the 'slow' datasets. SD_t after filtering was computed and corrected for the reduced number of degrees of freedom, and compared to SD_t of the unfiltered data. The average SD_t reduction over the 864 voxels with a significant CC-related artifact was 12.5 % for 'slow' (maximal improvement found was 55.3 %) and 11.3 % for 'fast' (maximal improvement

65.1 %). For the 2046 voxels with a RC-related artifact, the average SD_t improvement was 9.2 % for 'slow' (up to 52.7 % improvement) and 7.3 % for 'fast' (43.1 % maximum). Data filtered for both cardiac and respiratory artifacts showed an average improvement of 13.1 % for 'slow' and 12.0 % for 'fast' in the 271 voxels that suffered from both artifacts.

Performance of the method described here was compared to RETROICOR (Glover 2000). All data filtered with RETROICOR were otherwise processed in the same way, as described earlier. Filtering performance for all 6 volunteers is presented in the Figures 5B and 6B, using the same voxels shown in Figures 5A and 6A. The results obtained with the novel method are very similar to RETROICOR, albeit marginally better. There is a high degree of correlation between the data in Figures 5A and 5B, and Figure 6A and 6B, respectively. Pixels that perform poorly in the method described here typically also perform poorly in RETROICOR, further supporting the notion that the remaining spectral intensity is in all likelihood unrelated to CC or RC. Filtering using RETROICOR resulted in an 8.2 % improvement for 'slow' and 10.8 % for 'fast' when filtering for both cardiac and respiratory artifacts (averaged over the same 271 voxels used earlier). When only cardiac artifacts were filtered, the average SD_t reduction was 6.3 % for 'slow' and 9.2 % for 'fast'. Respiratory-only filtering resulted in an average 8.0 % and 6.2 % decrease in SD_t for respectively 'slow' and 'fast'. Note that in this RETROICOR implementation only first and second order terms were removed (as was described in (Glover 2000)), and the 2^{nd} and 3^{rd} harmonic of the cardiac artifact were therefore not removed. Even though the amplitude of these harmonics is only a fraction of the primary peak, they are distinguishable in the frequency spectrum of the voxel shown in Figure 3. Therefore, these harmonics could contribute to the residual artifact after RETROICOR filtering, which would hinder RETROICOR performance. This is the exact RETROICOR implementation described by Glover et al. (Glover 2000), and no attempt was made to optimize it.

One key difference between the method proposed here and RETROICOR is that RETROICOR computes the phase of the cardiac and respiratory cycle at which each data point was acquired, thereby assuming that the artifacts scale with the cardiac and respiratory interval. This approach is not necessarily correct, e.g. during the cardiac cycle a contraction of the heart muscle is followed by a period of relative rest. If the heart rate changes it is predominantly the length of this interval that changes. Even the respiratory cycle, which is more sinusoidal in character, shows periods of relative rest in between breathing action. The method described here therefore looks at the MR acquisition time relative to the nearest physiological event for averaging purposes, so that the artifact characteristics for each voxel can be determined. Scaling the bin size with changes in heart or respiratory rate would result in blurring of the artifact estimate. The acquisition time of the MR data relative to the nearest event was computed instead of using the nearest preceding event because the latency between the heart beat and its detection is unknown. It would be a trivial modification however to compute the correction relative to the nearest preceding event if so desired, for example if the heart rate is measured using ECG (in which case there is no significant detection latency). Ideally, the true latency should be determined and the window should be accordingly placed asymmetrically instead of centered on the time of detection of the artifact event.

Note that the SD_t reduction percentages found for 'slow' and 'fast' are not directly comparable, since the severity of the artifact in a given voxel is significantly affected by both flip angle and TR, which are substantially different for these two scans. Also, the relative contribution of intrinsic (image) noise to the overall temporal stability, as well as the degree of temporal correlation, is expected to differ for 'fast' and 'slow'. As was discussed in *Filtering Method*, a limited number of time points was not corrected. Volumes remained uncorrected if the acquisition time of the MRI volume was outside of the correction range, or if the volume belonged to a bin that contained less than 4 MRI time points. If the number of uncorrected time points is significant, it will reduce the performance of the filter. However, on average the

percentage of time points that was uncorrected was 1.3 % and 1.4 % for CC- and RC-filtering of 'slow', respectively, and less than 0.2 % for both CC- and RC-filtering of 'fast'. In RETROICOR all volumes were corrected. Possibly an intermediate method could be implemented, where modeling or fitting is applied to the artifact estimate. This could have performance benefits, e.g. allowing the correction of data assigned to under-populated bins, which remain uncorrected in the current approach, and would diminish the reduction of the number of degrees of freedom resulting from filtering.

Effect of Filtering on fMRI Performance

Applying the filter proposed here, as well as any other filtering method, affects the number of degrees of freedom, and therefore the performance of further analysis of the same data. Reducing the number of degrees of freedom will affect the noise distribution and therefore reduce the t-score. This penalty is typically low if the number of samples is sufficient. For example, for a p-threshold of 0.05 the loss will be less than 1 % if there are more than 85 degrees of freedom remaining.

There is however a more serious potential confound. If the cardiac and/or respiratory artifact occurrences (events) are not equally distributed between the 'active' and 'rest' state of the fMRI experiment there will be some degree of correlation between the fMRI analysis and the filtering. This correlation is typically small, but not necessarily non-zero. Note that this correlation can be the result of some physiological effect (e.g. arousal) or a truly random effect. If filtering is performed before fMRI analysis, the filter will remove a fraction of the BOLD signal. This is demonstrated by Figure 7, which shows the result of the Monte Carlo simulations. The broken line shows the fraction of the BOLD signal change that is measured as a function of the number of degrees of freedom used by the filter. The plot shows that the amount of activation that is detected is consistently below the true BOLD-related signal change when the filter is applied before fMRI analysis (broken line). This effect was found to be similar for all heart rates (results not shown), so the curves plotted here are averaged over the different heart rates used in the simulation. The measured activation amplitude is reduced for an increasing number of degrees of freedom used by the filter.

For a given number of degrees of freedom used, this effect is independent of the actual filtering algorithm used. Therefore, a similar penalty was found for RETROICOR, which uses 4 degrees of freedom in cardiac-only mode, shown by the symbol × in Figure 7. On the other hand, measuring activation amplitude before filtering would similarly introduce a bias in the fMRI results, so this is not a valid alternative. The correct approach is to apply the filter in the form of additional regressors, simultaneous with BOLD signal change amplitude detection, as was described as the 'simultaneous approach' in the *MATERIALS AND METHODS* subsection *Effect of Filtering on fMRI*. This eliminates the bias, as is shown by the solid line in Figure 7, which demonstrates that the true fMRI signal amplitude is measured, independent of the number of bins used for cardiac and/or respiratory filtering.

Analysis of the fMRI datasets resulted in 8152 selected voxels for all 6 volunteers combined. Figure 8 shows a plot of the t-score found in these voxels using the filtered design matrix as a function of the t-score in the same voxels found using the unfiltered design matrix. Data show that removal of physiological noise on average improved the t-score. The average improvement in t-score was 4.6 % and 4.0 % more voxels met the (arbitrary) selection criterion ($t \geq 5$).

CONCLUSION

A model-free, flexible, adaptive filter for MRI time series data has been demonstrated in an fMRI-like setting. In the majority of voxels it provides artifact suppression to noise level with no significant effect on signals unrelated to the cardiac and respiratory cycle, even when these

artifacts are undersampled in the MRI data acquisition. Under the current experimental conditions and for the level of variation in CC and RC encountered during these experiments the performance of this method is at least equivalent to the RETROICOR method. It has also been demonstrated that filtering should not be applied independently of fMRI activity detection, but concurrently, to avoid systematic underestimation of the BOLD-related signal change.

Acknowledgements

The authors thank Susan O'Flahavan (Advanced MRI, LFMI, NINDS, NIH) for her help with the experiments and Judy Ireland (LFMI, NINDS, NIH) for making the arrangements that allowed Roel Deckers to spend some time at NIH. This research was supported by the Intramural Research Program of the National Institute of Neurological Disorders and Stroke, National Institutes of Health. Gary Glover was supported by the National Center for Research Resources (NIH), grant number P41 09784.

References

Biswal B, DeYoe AE, Hyde JS. Reduction of physiological fluctuations in fMRI using digital filters. Magn Reson Med 1996;35:107–13. [PubMed: 8771028]

Biswal B, Yetkin FZ, Haughton VM, Hyde JS. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. Magn Reson Med 1995;34:537–41. [PubMed: 8524021]

Bodurka J, Ledden PJ, van Gelderen P, Chu R, de Zwart JA, Morris D, Duyn JH. Scalable multichannel MRI data acquisition system. Magn Reson Med 2004;51:165–71. [PubMed: 14705057]

Buonocore MH, Maddock RJ. Noise suppression digital filter for functional magnetic resonance imaging based on image reference data. Magn Reson Med 1997;38:456–69. [PubMed: 9339447]

Dagli MS, Ingeholm JE, Haxby JV. Localization of cardiac-induced signal change in fMRI. Neuroimage 1999;9:407–15. [PubMed: 10191169]

Dale AM, Buckner RL. Selective averaging of rapidly presented individual trials using fMRI. Hum Brain Mapp 1997;5:329–40.

de Zwart JA, Ledden PJ, van Gelderen P, Bodurka J, Chu R, Duyn JH. Signal-to-noise ratio and parallel imaging performance of a 16-channel receive-only brain coil array at 3.0 Tesla. Magn Reson Med 2004;51:22–6. [PubMed: 14705041]

de Zwart JA, van Gelderen P, Kellman P, Duyn JH. Application of sensitivity-encoded echo-planar imaging for blood oxygen level-dependent functional brain imaging. Magn Reson Med 2002;48:1011–20. [PubMed: 12465111]

Durbin J, Watson GS. Testing for serial correlation in least squares regression. III. Biometrika 1971;58:1–19.

Glover GH, Li TQ, Ress D. Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. Magn Reson Med 2000;44:162–7. [PubMed: 10893535]

Hu X, Kim SG. Reduction of signal fluctuation in functional MRI using navigator echoes. Magn Reson Med 1994;31:495–503. [PubMed: 8015402]

Hu X, Le TH, Parrish T, Erhard P. Retrospective estimation and correction of physiological fluctuation in functional MRI. Magn Reson Med 1995;34:201–12. [PubMed: 7476079]

Ishihara Y, Calderon A, Watanabe H, Okamoto K, Suzuki Y, Kuroda K, Suzuki Y. A precise and fast temperature mapping using water proton chemical shift. Magn Reson Med 1995;34:814–23. [PubMed: 8598808]

Lund TE, Madsen KH, Sidaros K, Luo WL, Nichols TE. Non-white noise in fMRI: does modelling have an impact? Neuroimage 2006;29:54–66. [PubMed: 16099175]

Ogawa S, Tank DW, Menon R, Ellermann JM, Kim SG, Merkle H, Ugurbil K. Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with magnetic resonance imaging. Proc Natl Acad Sci U S A 1992;89:5951–5. [PubMed: 1631079]

Pruessmann KP, Weiger M, Scheidegger MB, Boesiger P. SENSE: sensitivity encoding for fast MRI. Magn Reson Med 1999;42:952–62. [PubMed: 10542355]

Sakuma H, Fujita N, Foo TK, Caputo GR, Nelson SJ, Hartiala J, Shimakawa A, Higgins CB. Evaluation of left ventricular volume and mass with breath-hold cine MR imaging. Radiology 1993;188:377–80. [PubMed: 8327681]

Thévenaz, P.; Ruttiman, UE.; Unser, M. Iterative multi-scale registration without landmarks. IEEE international conference on image processing. IEEE international conference on image processing; Washington, DC, USA: 1995. p. 228-31.

Waldvogel D, van Gelderen P, Muellbacher W, Ziemann U, Immisch I, Hallett M. The relative metabolic demand of inhibition and excitation. Nature 2000;406:995–8. [PubMed: 10984053]

Watson GS. Serial correlation in regression analysis. I. Biometrika 1955;42:327–341.

Windischberger C, Langenberger H, Sycha T, Tschernko EM, Fuchsjager-Mayerl G, Schmetterer L, Moser E. On the origin of respiratory artifacts in BOLD-EPI of the human brain. Magn Reson Imaging 2002;20:575–82. [PubMed: 12467863]

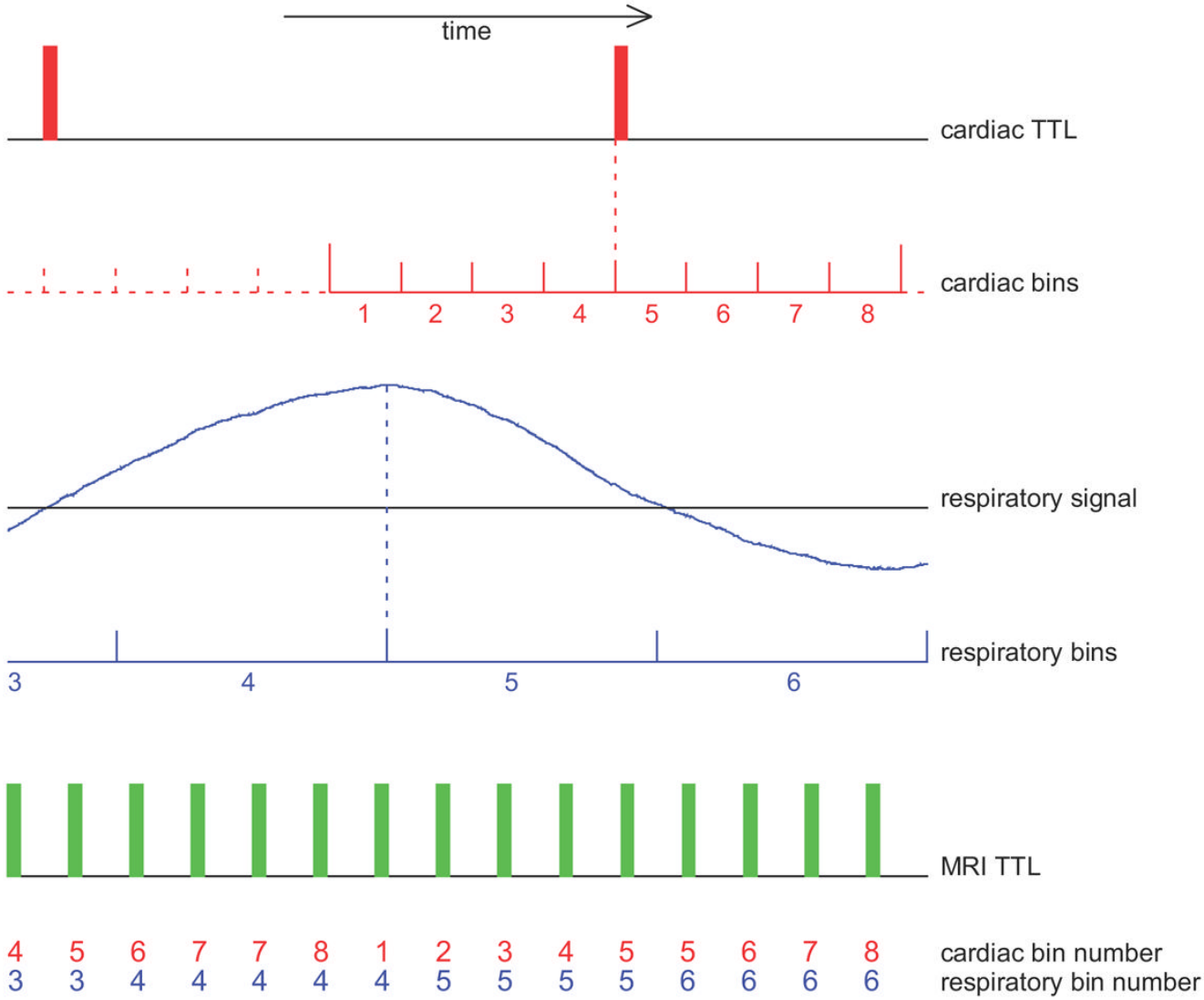

Figure 1.

Schematic representation of the filtering method. Both the cardiac cycle (CC) and respiratory cycle (RC) are subdivided into bins. In the example shown here, 8 bins are used for both cardiac and respiratory filtering. Information about CC and RC timing is derived from physiological monitoring data, acquired using a pulse oximeter and respiratory bellow, respectively. Filtering for CC and RC is performed independently. MRI data are assigned to a bin based on their acquisition time relative to the nearest event.

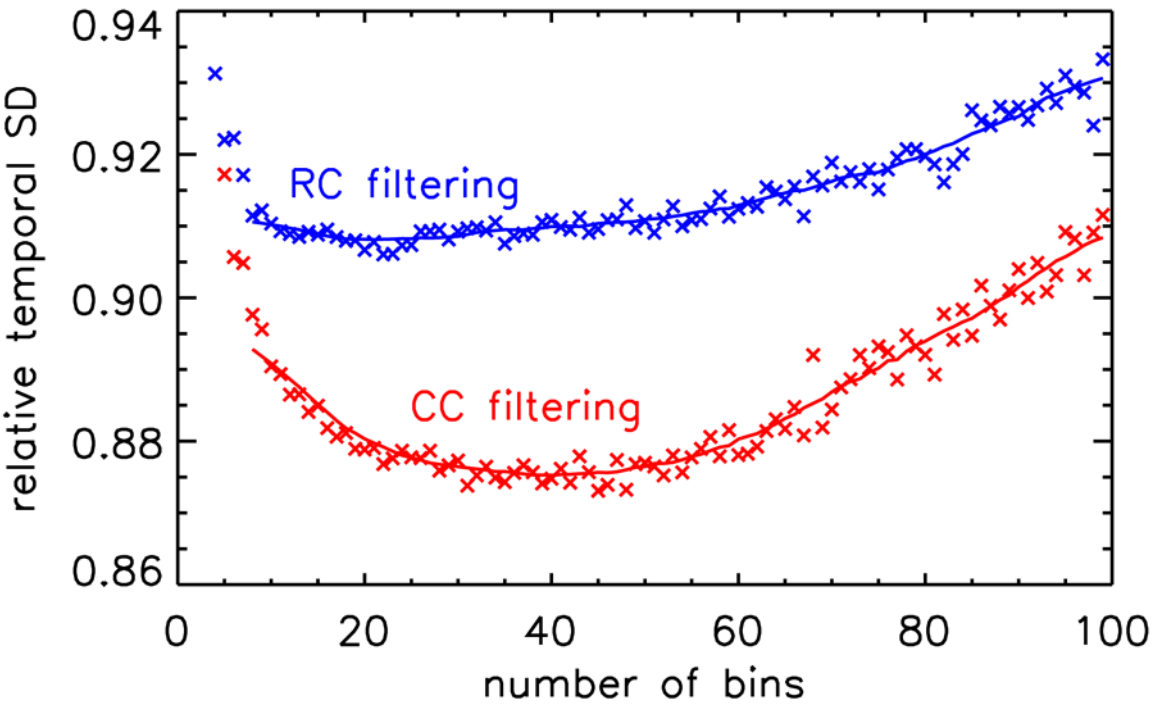

Figure 2.

Plot of the relative temporal standard deviation (SD_t) as a function of the number of bins used for filtering of the 'slow' data. The red symbols represent the result obtained for filtering cardiac artifacts, the blue symbols and line the result for filtering of respiratory artifacts. The average for all volunteers over all voxels suffering from a substantial artifact is shown (864 voxels for cardiac and 2046 for respiratory in total, see results). Symbols are the computed relative SD_t values as a function of the number of bins, the solid line is the result of smoothing these data 16-fold. Based on these data, 40 bins were used for filtering cardiac artifacts, and 21 for filtering respiratory artifacts.

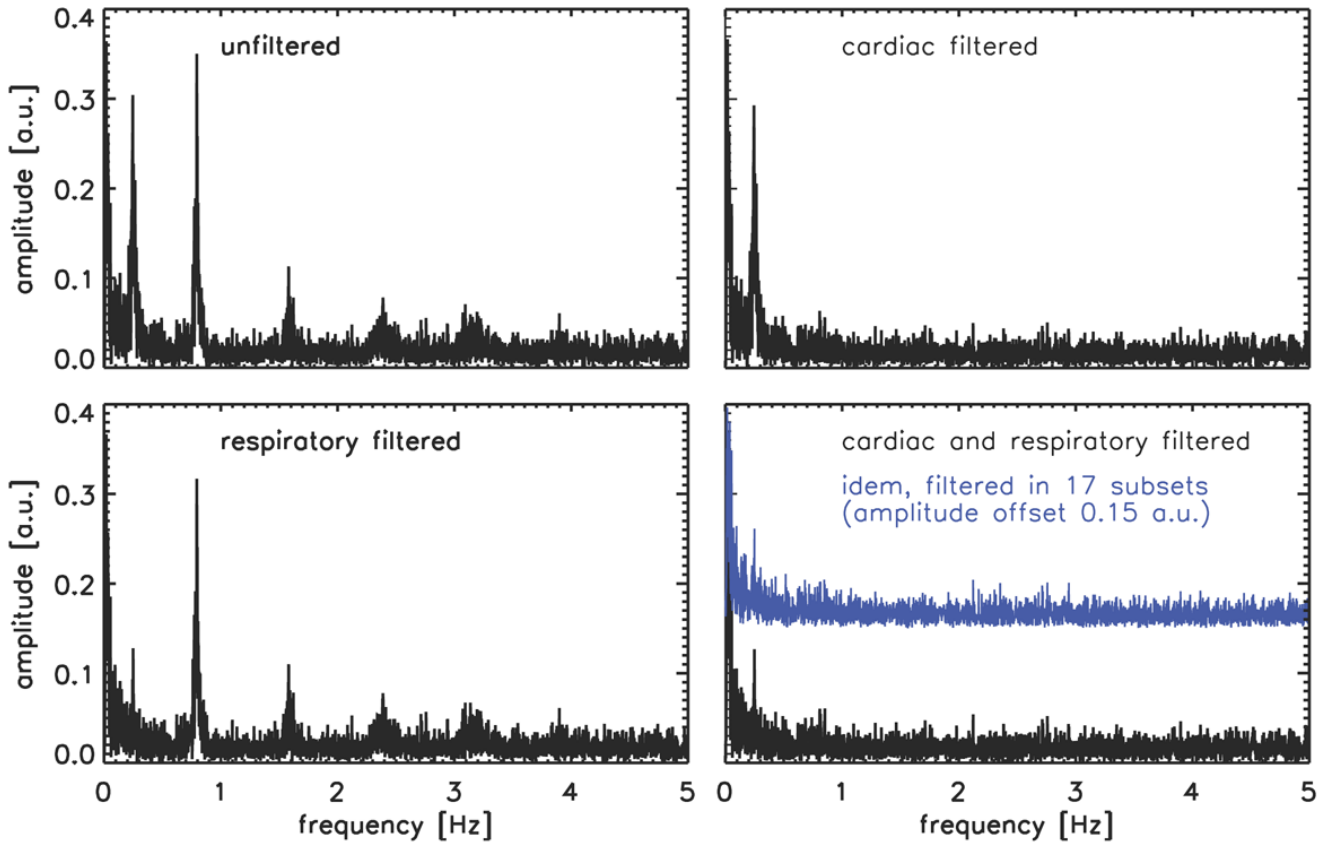

Figure 3.

Single-pixel frequency spectra for a 'fast' scan, obtained before and after cardiac or respiratory filtering, or both. The top-left plot shows the frequency spectrum for a pixel in the unfiltered dataset. The top-right plot shows the spectrum for the same pixel after application of the cardiac filter only. The lower-left plot shows similar data for respiratory-only filtered data, whereas the lower-right plot shows the frequency spectrum after application of both filters. The blue line in the lower-right plot, offset 0.15 units along the y-axis for figure clarity, is the result of filtering for cardiac and respiratory artifacts after splitting the data in 17 subsets (see text for details).

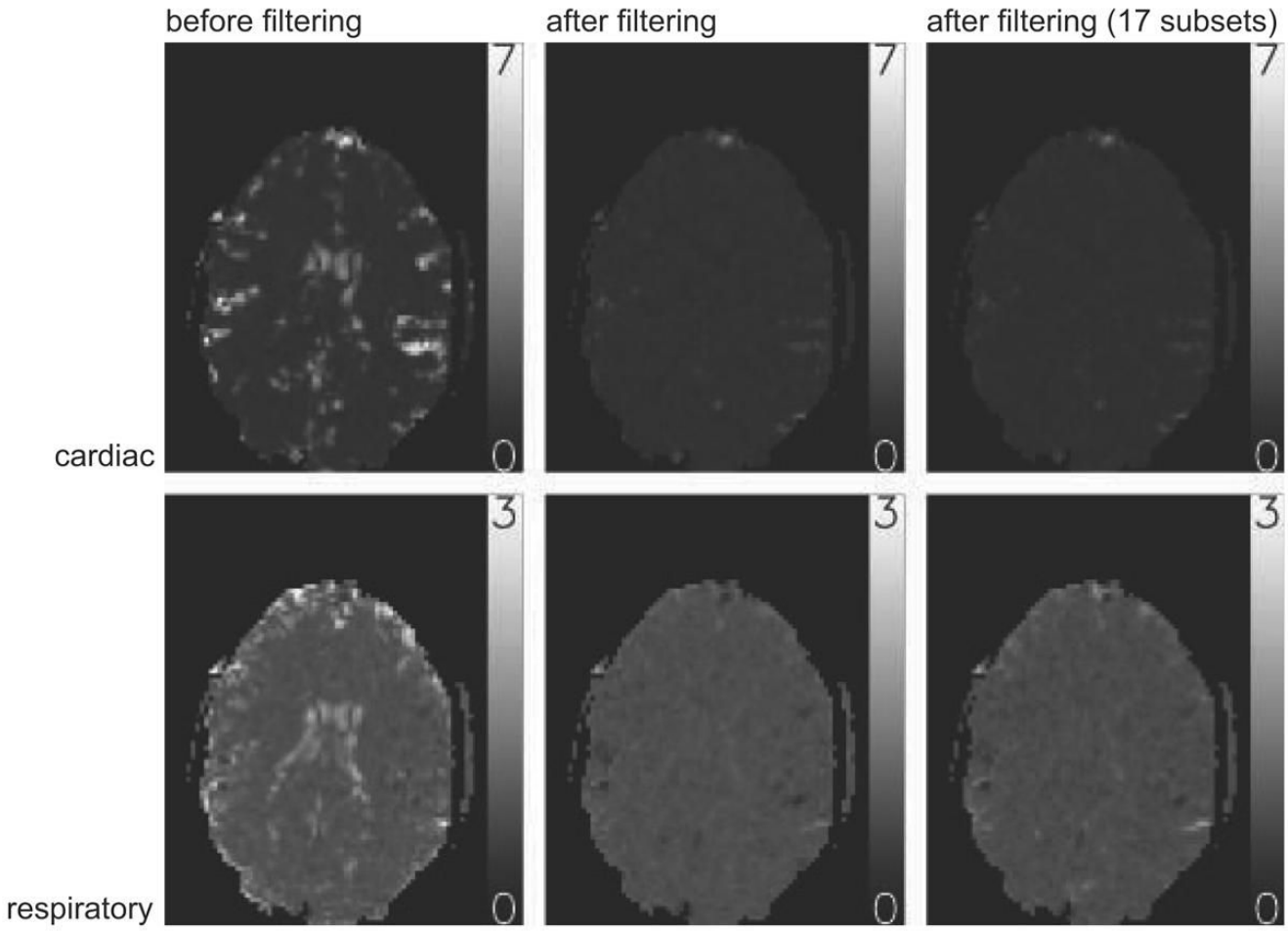

Figure 4.

Maps of the artifact intensity before and after filtering for the 'fast' scan from one of the volunteers. The maps show the average intensity in a frequency band surrounding the principal artifact frequency in the magnitude spectrum before (left column) and after (center and right column) filtering. The top row shows the intensity, expressed in units of baseline standard deviation, for the cardiac frequency range (1.18–1.48 Hz for this volunteer), the bottom row for the respiratory frequency range (0.20–0.50 Hz for this volunteer). The center column shows the artifact level map after filtering of all data simultaneously, the right column the result obtained when analyzing the data in 17 undersampled sets (see text for details).

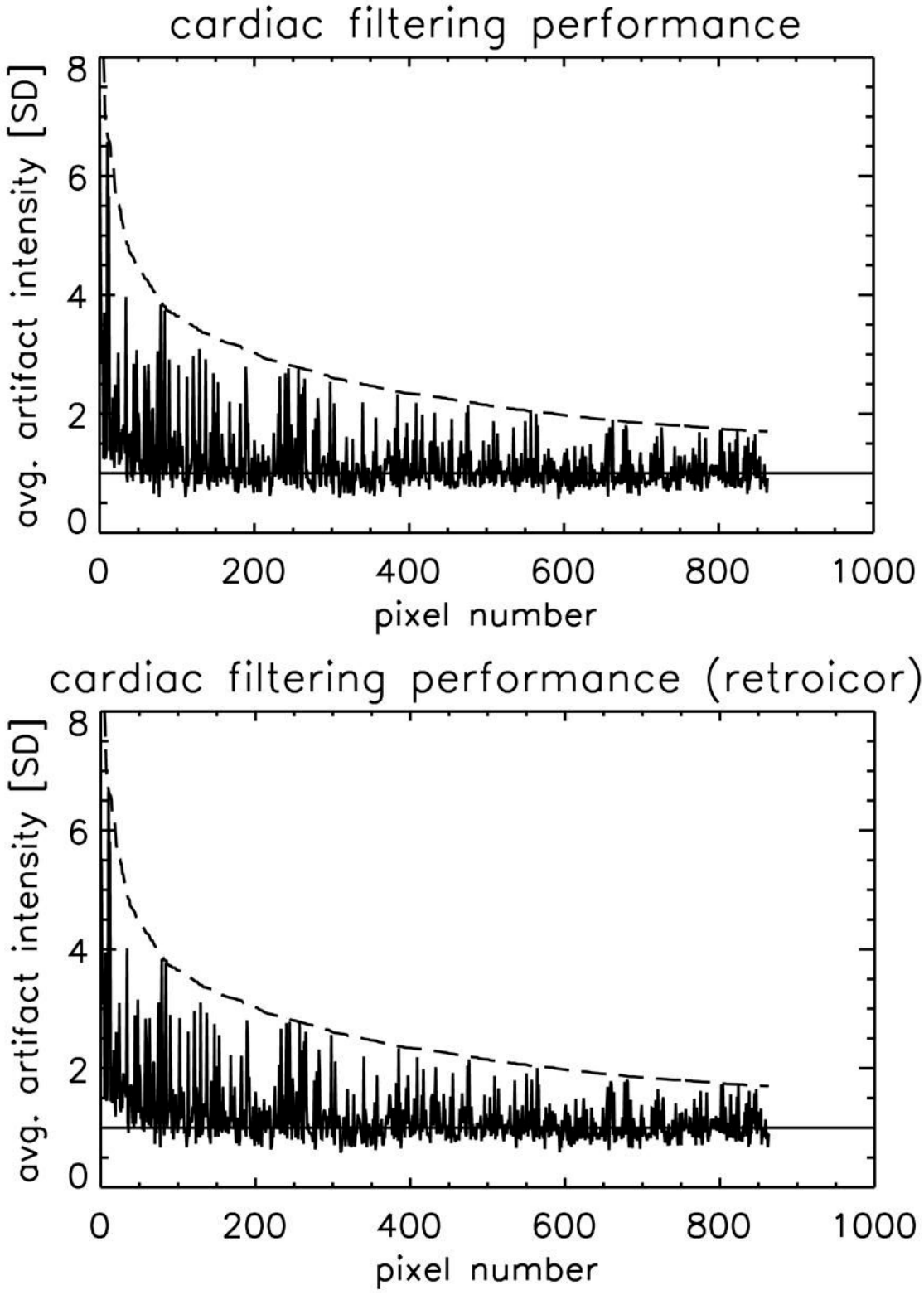

Figure 5.

Plots of the achieved artifact reduction for all voxels with a significant cardiac artifact in all of the 6 volunteers. Data are based on the decrease in average spectral intensity in a 0.15 Hz

band encompassing the principal cardiac artifact frequency. The dotted line represents the artifact level before filtering. It is a decreasing curve since voxels were sorted on pre-filtering artifact intensity. The horizontal line at 1 SD indicates noise level. The solid line shows the artifact intensity in the same frequency band after filtering. Figure 5A shows the result of filtering with the method described here, Figure 5B the result obtained using the RETROICOR method (Glover 2000).

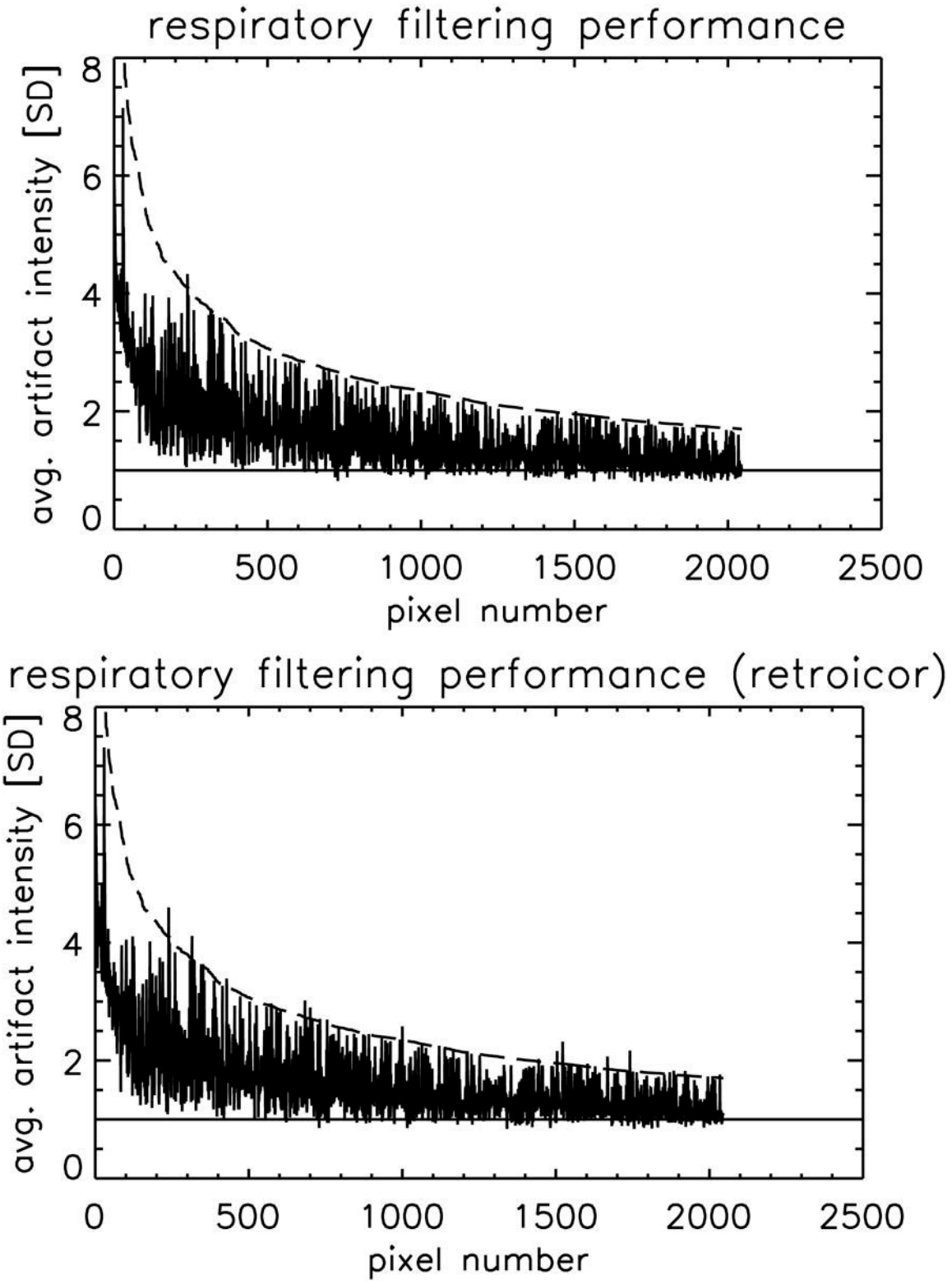

Figure 6.

Data similar to Figure 5, showing filtering performance for voxels with a significant respiratory artifact. Figure 6A shows the result obtained with the filter described here, Figure 6B the result obtained with the RETROICOR method (Glover 2000).

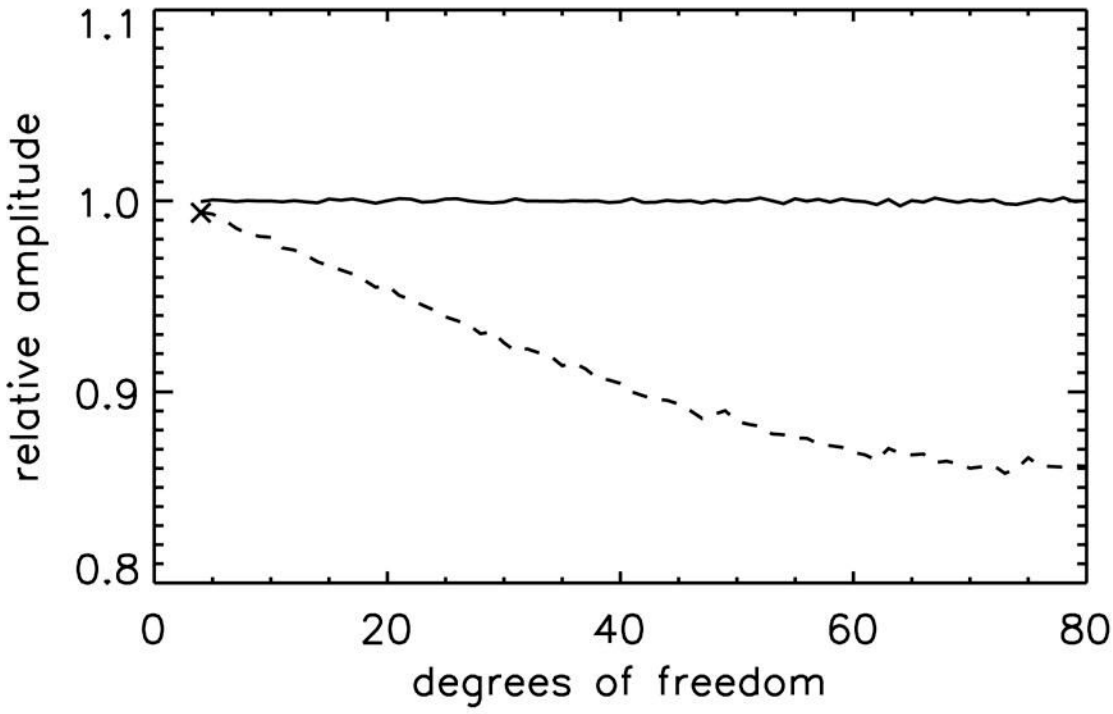

Figure 7.

Plot of the relative measured activation amplitude as a function of the number of degrees of freedom used by the filtering algorithm. These data are the average result of simulations for various cardiac rates, ranging from 40 to 120 beats-per-minute. Results demonstrate that when filtering is applied independently from and preceding the fMRI analysis, the measured activation amplitude is lower than the actual activation amplitude (dashed line). This is inherent to preprocessing, independent of the algorithm. RETROICOR shows a similar effect (×). This can be overcome by combining the filter and fMRI analysis into a single set of regressors for concurrent processing (solid line). Note that these plots show the detected percentage BOLD change, not significance. An increase in the number of degrees of freedom used by the filtering algorithm will negatively affect the t-score if the number of samples (acquired time points) is small.

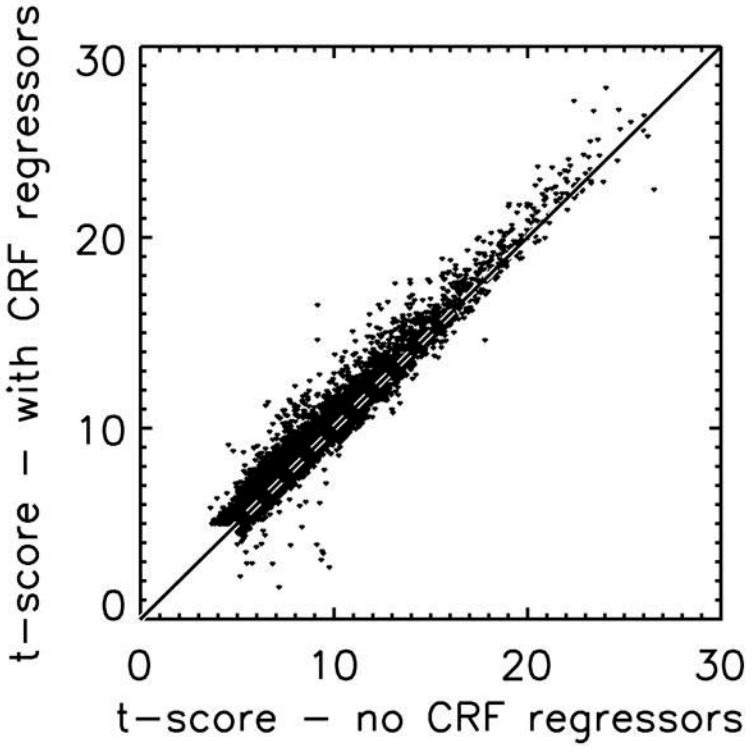

Figure 8.

Plot of the t-score in 8152 voxels, derived from functional data acquired on 6 volunteers. The x-axis shows the t-score computed when not accounting for cardiac and respiratory noise (no cardiac and respiratory filter (CRF) regressors), whereas the y-axis shows the t-score in the same voxels when analysis accounted for physiological noise through the inclusion of regressors for cardio- and respiratory artifacts in the design matrix. The solid line, encompassed by white dashes to ensure visibility in densely populated regions of the plot, represents equivalence (y = x).

Table 1

Overview of the observed heart and respiration rates in the various experiments. The length of the cardiac and respiratory bins (in ms) used during analysis is also shown. Note that 40 bins were used for cardiac filtering and 21 for respiratory filtering.

Volunteer		Avg. heart rate [s^{-1}] (SD)	Cardiac bin duration [ms]	Avg. respiration rate [s^{-1}] (SD)	Respiration bin duration [ms]
1	slow	1.27 (0.06)	23	0.33 (0.03)	180
	fast	1.33 (0.05)	21	0.35 (0.03)	169
2	slow	1.06 (0.05)	27	0.34 (0.03)	187
	fast	1.08 (0.06)	27	0.33 (0.03)	186
3	slow	0.82 (0.07)	34	0.24 (0.04)	294
	fast	0.80 (0.04)	34	0.25 (0.03)	260
4	slow	0.98 (0.10)	33	0.34 (0.04)	183
	fast	0.93 (0.12)	35	0.29 (0.03)	217
5	slow	0.91 (0.07)	32	0.28 (0.03)	205
	fast	0.94 (0.07)	32	0.28 (0.02)	204
6	slow	0.99 (0.07)	31	0.27 (0.02)	216
	fast	1.02 (0.08)	30	0.27 (0.03)	225