Evaluation of DISORDER: Retrospective Image Motion Correction for Volumetric Brain MRI in a Pediatric Setting


AJNR Am J Neuroradiol 2021, 42 (4) 774-781
doi: https://doi.org/10.3174/ajnr.A7001
http://www.ajnr.org/content/42/4/774
Evaluation of DISORDER: Retrospective Image Motion Correction for Volumetric Brain MRI in a Pediatric Setting


ABSTRACT

BACKGROUND AND PURPOSE: Head motion causes image degradation in brain MR imaging examinations, negatively impacting image quality, especially in pediatric populations. Here, we used a retrospective motion correction technique in children and assessed image quality improvement for 3D MR imaging acquisitions.

MATERIALS AND METHODS: We prospectively acquired brain MR imaging at 3T using 3D sequences, T1-weighted MPRAGE, T2-weighted TSE, and FLAIR in 32 unsedated children, including 7 with epilepsy (age range, 2–18 years). We implemented a novel motion correction technique through a modification of k-space data acquisition: Distributed and Incoherent Sample Orders for Reconstruction Deblurring by using Encoding Redundancy (DISORDER). For each participant and technique, we obtained 3 reconstructions as acquired (AQ), after DISORDER motion correction (Di), and with additional outlier rejection (DiOut). We analyzed 288 images quantitatively, measuring 2 objective no-reference image quality metrics: gradient entropy (GE) and MPRAGE white matter (WM) homogeneity. As a qualitative metric, we presented blinded and randomized images to 2 expert neuroradiologists who scored them for clinical readability.

RESULTS: Both image quality metrics improved after motion correction for all modalities, and improvement correlated with the amount of intrascan motion. Neuroradiologists also considered the motion corrected images as of higher quality (Wilcoxon z = −3.364 for MPRAGE; z = −2.066 for TSE; z = −2.645 for FLAIR; all P < .05).

CONCLUSIONS: Retrospective image motion correction with DISORDER increased image quality both from an objective and qualitative perspective. In 75% of sessions, at least 1 sequence was improved by this approach, indicating the benefit of this technique in unsedated children for both clinical and research environments.

ABBREVIATIONS: DISORDER = Distributed and Incoherent Sample Orders for Reconstruction Deblurring by using Encoding Redundancy; AQ = acquired; Di = after DISORDER motion correction; DiOut = Di with additional outlier rejection; GE = gradient entropy

Head motion is a common cause of image degradation in brain MR imaging. Motion artfacts negatively impact MR image quality and therefore radiologists’ capacity to read the images, ultimately affecting patient clinical care.1 Motion artfacts are more common in noncompliant patients,2 but even in compliant adults, intrascan movement is reported in at least 10% of cases.3 For children who require high-resolution MR images, obtaining optimal image quality can be challenging, owing to the

Received September 14, 2020; accepted after revision November 2.

Contribution of this study was supported in part by the National Institute for Health Research Mental Health Biomedical Research Center at South London, Maudsley NHS Foundation Trust, King’s College London, the National Institute for Health Research Mental Health Biomedical Research Center at Guys, and St Thomas’ Hospitals NHS Foundation Trust. Paper previously presented as a poster at: European Congress of Magnetic Resonance in Neuropediatrics, February 26-29, 2020, in Marseille, France.

Please address correspondence to Dr. Katy Vecchiato, MD, Centre for the Developing Brain, School Biomedical Engineering and Imaging Sciences, Department of Perinatal Imaging and Health, King’s College London, First Floor, South Wing, St Thomas’ Hospital, SE17EH, London, United Kingdom; e-mail: katy.vecchiato@kcl.ac.uk

Indicates open access to non-subscribers at www.ajnr.org

http://dx.doi.org/10.374/ajnr.21001

774 Vecchiato Apr 2021 www.ajnr.org
Here, we use a more general retrospective motion correction technique: Distributed and Incoherent Sample Orders for Reconstruction Deblurring by using Encoding Redundancy (DISORDER). In this method, k-space samples are reordered to enable retrospective motion correction during image reconstruction. Our hypothesis is that DISORDER improves clinical MR imaging quality and readability. To assess its use for clinical sequences, we acquired a dedicated epilepsy MR imaging protocol in 32 children across a wide age range. We used both objective image quality metrics and expert neuroradiologist ratings to evaluate the outcome after motion correction.

MATERIALS AND METHODS

Study Population

We recruited families for a prospective study of pediatric epilepsy (ethics ref 18/LO/1766). Informed consent was obtained from all participants or their parents, as appropriate. From June to November 2019, we recruited 32 participants: 25 healthy control participants and 7 children with focal epilepsy, ages 2–18 years (median 11), including 16 females (50%) (Table). Exclusion criteria were age younger than 6 months or older than 18 years, major neurologic conditions unrelated to epilepsy, and contraindications for 3T MR imaging.

Image Acquisition

Children were scanned without sedation on a 3T Achieva-TX (Philips Healthcare) using a 32-channel head coil. They were asked to stay still during scanning while watching a movie. The protocol was T1-weighted MPRAGE: TR = 7.7 ms, TE = 3.6 ms, flip angle = 8°, TI = 900 ms, echo-train length = 154 and acquisition time = 286 s; T2-weighted: TSE: TR = 2500 ms, TE = 344 ms, echo-train length = 133, acquisition time = 342 s; T2-weighted FLAIR: TR = 5000 ms, TE = 422 ms, TI = 1800 ms, echo-train length = 182, acquisition time = 510 s. Parallel imaging acceleration (SENSE) of 1.4 was used along both phase-encoding directions. Field of view was 240 × 188 × 240 mm, and images were 1 mm isotropic. The combined acquisition time was approximately 22 minutes.

All scans were acquired using the DISORDER scheme (Fig 1). A shot of k-space is defined as a portion of k-space phase-encoding data in the k1k2 plane, acquired within a single acquisition block. In Fig 1, each shot is represented by a different color. As demonstrated by Cordero-Grande et al. DISORDER aims to improve motion tolerance by guaranteeing that the acquisition of every shot contains a series of samples distributed incoherently throughout k-space. This is achieved with a modified phase-encoding sampling order. We adopted the “random-checkered” approach illustrated in Fig 1. Data are acquired in the inferior-superior k1, anteroposterior k2, and left–right k3 orientations; this way rotations on the sagittal plane (k1k2) are sampled faster within each shot, improving robustness to intrashot

---

**FIG 1.** The different k-space data acquisitions. On the left side, the standard acquisition that sequentially acquires adjacent lines in the grid, with an example shot given as bigger dots in blue. The image on the right represents DISORDER’s “random-checkered” acquisition, in which every shot acquires distributed information in k-space with a certain degree of randomness.

---

**TABLE**

Descriptive demographics of the study population

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Healthy Control Participants</th>
<th>Patients with Epilepsy</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>25</td>
<td>7</td>
<td>32</td>
</tr>
<tr>
<td>Age at scan (years)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>11.32 ± 4.8</td>
<td>11.6 ± 3.7</td>
<td>11.4 ± 4.5</td>
</tr>
<tr>
<td>1–5</td>
<td>3</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>6–10</td>
<td>9</td>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td>11–15</td>
<td>7</td>
<td>3</td>
<td>10</td>
</tr>
<tr>
<td>16–18</td>
<td>6</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>13</td>
<td>3</td>
<td>16</td>
</tr>
<tr>
<td>Female</td>
<td>12</td>
<td>4</td>
<td>16</td>
</tr>
</tbody>
</table>
motion. In our protocol, the numbers of shots are 120 (duration, 1200 ms each) for MPRAGE, 135 (658 ms) for TSE, and 100 (859 ms) for FLAIR.

There was not a substantial difference in image quality between data acquired via standard acquisition and DISORDER acquisition. This was tested in the pilot phase of the study for all structural images (MPRAGE, TSE, and FLAIR).

**Motion Correction**

Motion and reconstruction were estimated jointly using a parallel $k$-space model in the presence of rigid motion. Starting from a standard reconstruction assuming no motion, a first approximation of the motion parameters for each shot is obtained by maximizing the likelihood of the $k$-space measures for current reconstructed volume. Then a new volume is reconstructed with current motion parameters, and the method alternates between motion estimation and reconstruction until convergence. Motion correction is performed as part of the reconstruction stage; an in-depth description of the reconstruction algorithm has been described previously.

The time for reconstruction varied from 5 to 40 minutes, depending on degree of intrascan motion. In the case of high intrashot motion, the DISORDER framework can further improve the image quality by dismissing outlier shots. Therefore, each subject had their images reconstructed in 3 different ways: as acquired without motion correction (Aq), with DISORDER reconstruction (Di), and with DISORDER reconstruction including outlier shot rejection (DiOut).

Motion estimates from DISORDER correction can also provide a measure of intrascan motion. To summarize and quantify this intrascan motion, we averaged the temporal standard deviations of the 3 rotation parameters (in degrees) for every scan.

**Image Quality Assessment**

For all quality assessments, there were 288 images available (32 participants, 3 imaging modalities, and 3 reconstructions). To objectively compare image quality, we used 2 metrics that do not rely on reference datasets: gradient entropy (GE) and white matter (WM) signal homogeneity (MPRAGE only).

The entropy of an image is a measure of sharpness that characterizes its texture based on intensity. GE has been previously used to characterize image definition, smaller when areas of uniform signal intensity are separated by sharp edges. We calculated the normalized GE for Aq, Di, and DiOut. Decreased GE indicates that image information is concentrated at the edges, a measure of
sharpness. This metric has a high correspondence with visual assessment of clinical MR imaging.\textsuperscript{18}

WM signal homogeneities of the T1-weighted images were obtained using an automatic segmentation in FreeSurfer\textsuperscript{19} (version 6: http://surfer.nmr.mgh.harvard.edu). After calculating a WM mask, the mean and SD of the signal intensity were computed within the mask. The WM signal homogeneity was then measured as the mean scaled by SD, with higher WM signal homogeneity associated with higher image quality.

Image quality was further visually inspected by 2 pediatric neuroradiologists with more than 9 years’ experience each (A.E. and O.C.). Radiology scoring was explicitly for assessment of focal epilepsy, which needs very high contrast between gray matter (GM) and WM. They scored the images with a 4-point Likert scale: 1, unreadable (not suited for clinical use); 2, poor quality (main structures identifiable but heavily blurred or artifacts covering $\geq 50\%$ of the image); 3, good quality (good GM–WM differentiation, little blurring, or minor artifacts); and 4, excellent quality (no motion artifacts, good contrast, and perfectly defined GM–WM boundaries). Sagittal, coronal, and axial views of all images were presented to the 2 radiologists on the same screen and room environment in a randomized and blinded fashion. Each rater looked at all 288 cases in several sessions. The image viewer was rview (https://biomedia.doc.ic.ac.uk/software/irtk). In all reported comparisons (quantitative metrics and quality ranks), Wilcoxon signed-rank tests were used.

**RESULTS**

**Quantitative Metrics**

As expected, older children tended to move less than younger ones. This was a consistent relationship, statistically significant for MPRAGE and FLAIR (Spearman rho: $-0.416$, $-0.363$, both $P < .05$) but not for TSE scans (Spearman rho: $-0.229$, $P = .21$) (Fig 2).

GE was reduced after motion correction across all modalities: the Wilcoxon rank test performed on mean scores before and after motion correction showed a statistically significant difference (for Di, $z = -4.861$ MPRAGE, $z = -4.769$ TSE, $z = -4.884$ FLAIR; for DiOut, $z = -4.937$ MPRAGE, $z = -4.937$ TSE, $z = -4.938$ FLAIR; all $P < .05$) (Fig 3).

There was a linear association between GE decrease after motion correction and degree of intrascan motion. We calculated the difference in GE before and after motion corrected data and estimated a linear regression against motion for each technique. The coefficients of determination ($R^2$) for MPRAGE images against motion were $R^2 = 0.24$ in Aq-Di and $R^2 = 0.48$ in Aq-DiOut; for T2-weighted images, they were $R^2 = 0.63$ in Aq-Di and $R^2 = 0.69$ Aq-DiOut; for FLAIR images, they were $R^2 = 0.44$ in Aq-Di and $R^2 = 0.51$ in Aq-DiOut; all $P < .05$. The reduction of GE was larger after outlier rejection (Fig 4).

There was an increase of WM signal homogeneity after motion correction on the MPRAGE images ($R^2 = 0.16$ for Aq-Di and $R^2 = 0.15$ for Aq-DiOut; both $P < .05$). One case was
excluded from this analysis because of large motion during the MPRAGE acquisition and the resulting failure of the FreeSurfer pipeline for the uncorrected reconstruction. Full-brain segmentation was possible on the motion-corrected version of the images, and WM signal homogeneity was measured (Fig 5).

**Qualitative Metrics**

Expert visual inspection showed that image quality generally improved after motion correction (Fig 6). There was agreement between observers in raw image scores according to Cohen kappa coefficient for interrater reliability: kappa > 0.3 for FLAIR and MPRAGE images, kappa > 0.6 for TSE images, all P < .05. The intra-class correlation of the score change was used as another measure of interrater consistency on rating improvement. The intra-class correlation coefficient for absolute agreement in the change in scores after motion correction of the images was 0.8 for TSE (Di and DiOut); for FLAIR, it was 0.64 for Di and 0.56 for DiOut images; all P < .05. The rating increase was less consistent between observers for MPRAGE images: 0.52 (P < .05) for Di and 0.37 (P = .09) for DiOut.

Wilcoxon signed-rank tests were conducted to compare the expert scores before and after image correction: the improvement of scores was statistically significant both for Di and DiOut and for both radiologists (for observer 1, z = −3.164 MPRAGE, z = −2.066 TSE, z = −2.645 FLAIR; for observer 2, z = −3.162 MPRAGE, z = −2.714 TSE, z = −3.419 FLAIR; all P < .05).

As expected, the Spearman rho showed a negative correlation between the amount of motion and the resulting score of acquired images in all modalities and for both observers (observer 1: ρ = −0.558 for MPRAGE, ρ = −0.496 for T2, both P < .05, ρ = −0.216 for FLAIR, P = .24; observer 2: ρ = −0.619 for MPRAGE, ρ = −0.641 for T2, ρ = −0.544 for FLAIR, all P < .05).

**DISCUSSION**

Successful neuroimaging in children is important for both clinical evaluation and research in brain development and disease. However, obtaining high-quality data in children is challenging in the highly motion-sensitive MR imaging context. In this work, we demonstrated the benefits of retrospective motion correction.
correction on a non-sedated pediatric cohort undertaking brain MR imaging on a 3T scanner using the DISORDER framework. We applied this method to a dedicated high-resolution epilepsy protocol across a wide age range and showed that DISORDER motion correction increases image quality both quantitatively and qualitatively.

The 2 neuroradiologist raters broadly agreed on the improved diagnostic value of motion-corrected images; on average, ratings were higher after DISORDER. In 24/32 (75%) participants, at least 1 technique was improved by DISORDER in a clinically significant way—from being considered unreadable or of poor quality to good or excellent quality. Images acquired in the presence of no or little motion maintained their high quality after motion correction.

In a small number of cases, DISORDER reconstructions were rated lower, though all in the context of very low motion.

FIG 6. Scores before and after motion correction (DiOut) for all modalities. Red and blue correspond to observer 1 and 2’s scores, respectively. The dot indicates the score before motion correction, and the triangle indicates the score after correction. Motion correction generally improved the image quality from a radiologic perspective.
However, an advantage of this retrospective method is that images both before and after motion correction are always available for radiologic evaluation. Observed differences between raters are in line with previous studies in which differences in subjective radiologic judgment are reported.

As expected from practical experience and previous studies, younger children tended to move more than older ones, which was observed statistically significant for MPRAGE and FLAIR, though not for TSE. None of the participants exhibited very high motion during this acquisition. This was the third sequence acquired in our protocol; it may be that participants at that time point were simply more settled or comfortable (engaging with the movie they were watching or spontaneously falling asleep).

The proposed motion compensation method is particularly flexible for use, and it is applicable to any volumetrically encoded sequence. It does require modification of the scanner software to meet the requirements of data acquisition ordering, but it does not involve any additional hardware, relevant manipulation of imaging parameters, or additional operator training. The image reconstruction is operated with a vendor-independent off-line open-source code (https://github.com/mriphysics/DISORDER/tree/1.1.0), so the technique is not restricted to a specific vendor and has been tested on scanners from several manufacturers.

To comply with the requirements of enough SNR for high image resolution (1 mm) and strong motion tolerance, data are acquired with moderate acceleration factors (SENSE factor 1.4 × 1.4). However, the DISORDER encoding does not increase the sequence acquisition time per se. In practice, it may reduce the need for repeat scans and the time overhead would compare favorably with times required for sedation. The method can also provide motion correction for additional 3D sequences in which motion correction can be crucial (eg, relaxometry).

The approach provides clinically useful improvements. In this study, it is applied to a dedicated epilepsy protocol in which clinicians aim to identify sometime subtle abnormalities such as focal changes in cortical thickness, subcortical signal abnormalities, or blurring of the GM–WM junction. In the presence of motion, DISORDER can be a helpful tool because these image features can be enhanced as shown in Fig 7.

More broadly, DISORDER could be helpful in a clinical setting to improve identification of other types of lesions not necessarily related to epilepsy including smaller injuries such as punctate bleeds (in which small motion may blur out the injury) or more obvious such as a brain mass (in which the extent and edge of pathologic tissue can be difficult to discern on blurry images). This method would be beneficial not only for children but also for patients with high anesthesiologic risk, situations in which time constrains the possibility to repeat scans, and for adult patients with intellectual disabilities.

Some limitations are noted. First, the method may require longer scan times to perform well in cases of very quick, large range, or continuous motion, and this is not assessed here. Second, DISORDER sampling increases motion sensitivity, facilitating its subsequent correction, so some enhancement of artifact levels compared with standard acquisition schemes is likely on the uncorrected images. Two further clinical considerations are also not yet addressed, the impact of reconstruction delay (up to 1 hour) and a quantification of maximum tolerable degradation before an acquisition needs to be repeated (how bad an image can be before it needs to be repeated). However, given the almost global improvement in data quality for motion-corrupted data, the reconstruction delay is probably not a large concern and will be addressable with future software implementations. Certainly, in this case, no DISORDER-reconstructed sequences were considered radiologically unreadable.

This framework for motion-tolerant structural 3D brain images improves clinical MR imaging quality both quantitatively and qualitatively. This might have substantial safety and economic implications for health care, reducing the clinical indication for sedation and repeat scans in children and adults.

ACKNOWLEDGMENTS

We are grateful to the families who generously supported this study. We also thank the Pediatric Neurology team from the Evelina London Children Hospital, including Dr. Ruth Williams, Dr. Elaine Hughes, Dr. Shan Tang, and Dr. Karine Lascelles.

REFERENCES

1. Andre JB, Bresnahan BW, Mossa-Basha M, et al. Toward quantifying the prevalence, severity, and cost associated with patient...