



Motion Artefact Reduction in Abdominal Magnetic Resonance Imaging Using Deep Learning Reconstruction

Marko Olsen^{1*}, Stefan Farouk², Lerato Milosevic³

¹Department of Medical Sciences, Tamiraparani Center for Medical Sciences, Netherlands

²Department of Medical Sciences, Congo Institute of Medical Sciences, Thailand

³Department of Medical Sciences, Lapland Research Institute of Medical Sciences, Canada

Article Info

Volume 1, Issue 1, 2026

doi: [10.62715/ijrmit.2026.v1.i1.pp21-26](https://doi.org/10.62715/ijrmit.2026.v1.i1.pp21-26)

Received: 10 May 2026

Accepted: 21 July 2026

Published: 06 September 2026

Abstract

Motion is the commonest cause of non-diagnostic abdominal magnetic resonance imaging. Breathing, cardiac pulsation and peristalsis produce ghosting and blurring that conventional countermeasures, among them breath-holding, respiratory triggering, radial sampling and parallel imaging, reduce but do not remove. Deep learning reconstruction has been introduced on clinical scanners chiefly as a means of recovering image quality from undersampled or noisy data, and its relevance to motion is largely indirect: by permitting shorter acquisitions it shortens the breath-hold that has to be sustained. Networks that correct motion-corrupted data directly are an active research topic and are less mature. This article proposes a staged procedure by which an imaging department can introduce deep learning reconstruction into abdominal protocols with motion reduction as the explicit aim. The six stages cover definition of the clinical target, redesign of the acquisition, choice and configuration of the reconstruction, technical verification, a blinded reader comparison, and monitoring in routine use. Each stage is described with the reasoning behind it and the record it should leave. The procedure is assembled from the published literature and has not been applied or validated here. Its limits, in particular the risk that a learned reconstruction conceals or invents fine detail, are set out.

Keywords: Abdominal MRI, Motion artefact, Deep learning reconstruction, Breath-hold imaging, Image quality evaluation, Protocol optimisation

© 2026 Marko Olsen et al. This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

1. Introduction

Abdominal magnetic resonance imaging (MRI) asks for something difficult. Sequences that last many seconds must be acquired with the breath held, repeatedly, by people who are often elderly, breathless or in pain. When the breath-hold fails, the image is degraded by ghosting and blurring, and the examination may have to be repeated or reported with reservations. Motion artefact is a long-standing problem of the technique as a whole, and a standard review of the field describes it as one with many partial solutions and no complete one (Zaitsev et al., 2015).

* Corresponding author: Marko Olsen, Department of Medical Sciences, Tamiraparani Center for Medical Sciences, Netherlands. E-mail: marko.olsen@tamiraparani.edu

© 2026 Marko Olsen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Reconstruction based on deep learning is the most recent addition to the list. Within a few years it has moved from research papers to products offered by the major manufacturers, on the strength of evidence that trained networks can reconstruct diagnostic images from data too sparse or too noisy for conventional methods ([Hammernik et al., 2018](#); [Knoll et al., 2020](#)). Departments acquiring the software face a practical question that the technical literature does not answer for them: how to change abdominal protocols so that the new reconstruction reduces motion artefact, and how to be satisfied that nothing of diagnostic value is lost in the exchange.

This article proposes a procedure for doing so. It is a method paper: it describes stages, reasons and records, and reports no measurements, examinations or reader results. The procedure has not been validated, and the final sections say what validation would require.

2. Background

The design of the procedure depends on three things: how motion produces artefact, what the established countermeasures do, and what deep learning reconstruction does and does not change.

2.1. How motion corrupts the image

MRI data are collected in spatial-frequency space over time, and the reconstruction assumes that the object is the same throughout. When it is not, the inconsistencies between lines of data appear as artefact, whose form depends on the motion and on the order of sampling. Periodic motion during a Cartesian acquisition produces discrete ghosts displaced along the phase-encoding direction; irregular motion produces diffuse blurring and loss of edge sharpness ([Zaitsev et al., 2015](#)). Table 1 lists the sources that matter in the abdomen with their usual remedies. One deserves particular note. Transient breathlessness shortly after injection of gadoxetate disodium can spoil the arterial phase of a liver examination, the one phase that cannot be repeated ([Davenport et al., 2013](#)).

Table 1: Sources of motion in abdominal MRI, their usual remedies and the possible contribution of deep learning reconstruction			
Source of motion	Typical artefact	Conventional remedy	What deep learning reconstruction can add
Respiration, failed breath-hold	Ghosting along the phase-encoding direction; blurring	Shorter breath-hold, triggering, navigators	Shorter acquisition through higher acceleration
Respiration, free breathing	Streaks with radial sampling	Golden-angle radial with motion sorting (Feng et al., 2016)	Removal of undersampling streaks; faster reconstruction
Cardiac pulsation	Ghosts across the left hepatic lobe	Change of phase direction, saturation bands	Little direct effect
Peristalsis	Blurring of bowel wall and nearby organs	Antiperistaltic agent, single-shot sequences	Denoising of fast single-shot images
Breathlessness after contrast	Degraded arterial phase	Several short arterial phases, radial acquisition	Support for very short or continuous acquisitions
Bulk body movement	Global blurring, misregistration	Coaching, repeat scan, PROPELLER (Pipe, 1999)	Research-stage correction networks (Küstner et al., 2019)

2.2. Conventional countermeasures

Existing remedies work in one of three ways. Some stop the motion or wait for it: breath-holding, respiratory triggering and navigator gating. Some make the acquisition less sensitive to it: radial and PROPELLER trajectories sample the centre of k-space repeatedly, so that motion is averaged into mild blurring or streaks in place of coherent ghosts ([Pipe, 1999](#)). The rest make the acquisition faster. Parallel

imaging with SENSE or GRAPPA uses the spatial sensitivity of receiver coils to replace some phase-encoding steps (Pruessmann *et al.*, 1999; Griswold *et al.*, 2002), and compressed sensing exploits image sparsity to go further (Lustig *et al.*, 2007). Combined with golden-angle radial sampling, these gave rise to free-breathing dynamic imaging and to reconstructions that sort data by respiratory state (Feng *et al.*, 2014; Feng *et al.*, 2016). Each has a cost: noise amplification, residual aliasing, long computation or a longer examination.

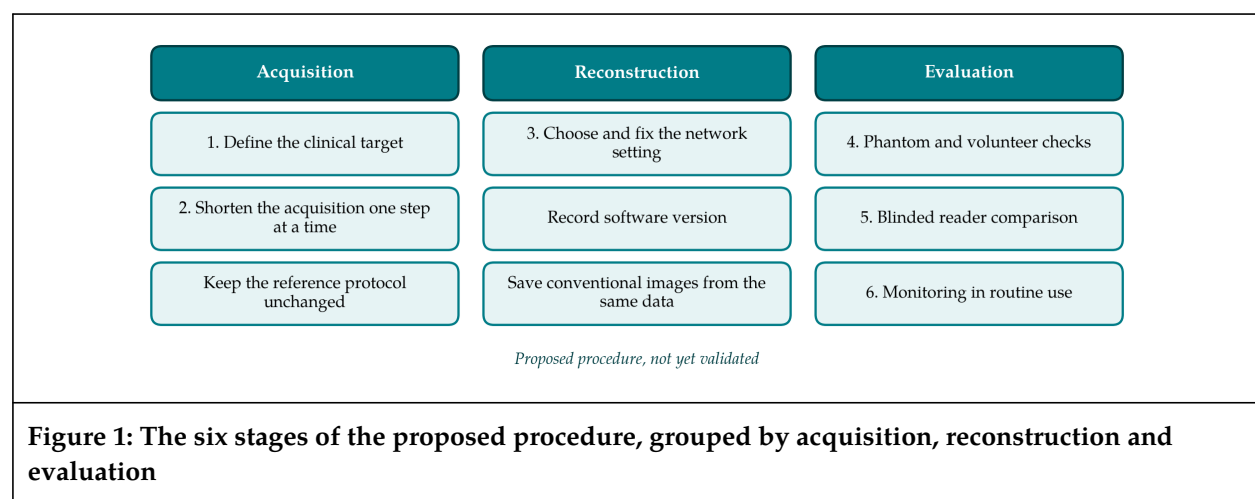
2.3. What deep learning reconstruction does

Most clinical deep learning reconstructions belong to one family. An iterative reconstruction is unrolled into a fixed number of steps, each alternating a data-consistency operation, which keeps the estimate faithful to the measured samples, with a learned regulariser in the form of a convolutional network. The variational network, the deep cascade and model-based deep learning are the reference designs (Hammernik *et al.*, 2018; Schlemper *et al.*, 2018; Aggarwal *et al.*, 2019). Trained on pairs of fully sampled and undersampled data, such networks suppress noise and aliasing at acceleration factors where parallel imaging fails. Their effect on motion is indirect. They do not model motion; they make it possible to acquire the same image in less time, and a shorter breath-hold is one that more people can manage.

Direct correction is a separate line of work. Image-to-image networks have been trained to map motion-corrupted images to clean ones, and hybrid schemes use a network to guide a physical model of the motion (Küstner *et al.*, 2019; Haskell *et al.*, 2019). These are promising and largely unavailable in clinical products. Both kinds share a known weakness. A learned prior can produce images that look sharp and plausible while departing from the data, and small structures or perturbations unlike anything in the training set can be removed or altered (Antun *et al.*, 2020). The best-known public training set consists of knee and brain examinations (Zbontar *et al.*, 2018), so abdominal performance cannot be assumed from published benchmarks.

3. The Proposed Procedure

The procedure has six stages, grouped in Figure 1 under acquisition, reconstruction and evaluation. It assumes a department with a scanner on which a deep learning reconstruction, supplied by the manufacturer or developed in research, is available for abdominal sequences, and a physicist or senior radiographer able to modify protocols.



3.1. Stage 1: Define the clinical target

The first decision is which motion problem is to be solved, for which patients and on which sequences. The answer should come from the experience the department has of failed examinations. Typical candidates are the breath-held T1-weighted gradient-echo sequence used for dynamic liver imaging, T2-weighted single-shot and fast spin-echo sequences, and diffusion-weighted imaging. A target is stated in operational terms, for example that the dynamic sequence should be acquired in a breath-hold short

enough for people who cannot hold for the present duration. A sequence with no identified motion problem is left alone.

3.2. Stage 2: Redesign the acquisition

Deep learning reconstruction reduces motion artefact only if the acquisition is changed to take advantage of it. The options are a higher parallel-imaging or compressed-sensing factor, fewer signal averages, or partial Fourier sampling, each of which shortens the scan and lowers the signal-to-noise ratio that the network is then asked to restore. Changes are made one at a time from the existing protocol, which is kept unaltered as the reference. Contrast-determining parameters, that is repetition time, echo time, flip angle and fat suppression, are left as they are, so that any change in lesion conspicuity can be attributed to the reconstruction and not to altered weighting. Where a radial free-breathing sequence is available, it is treated as a separate candidate and not mixed with the breath-hold variant (Feng et al., 2014).

3.3. Stage 3: Choose and configure the reconstruction

Commercial implementations expose few controls, usually a strength or denoising level and sometimes an edge-enhancement option. The setting is chosen on test data before any comparison and then fixed, because adjusting it case by case makes the result irreproducible. The software version, the setting and the sequences to which it is applied are written down; a version update is treated as a new reconstruction requiring at least the technical checks of the next stage. Throughout the evaluation the conventionally reconstructed images from the same raw data are also saved wherever the system allows. They cost nothing in scan time and provide the only direct check on whether a feature seen, or not seen, on the deep learning images is present in the data.

3.4. Stage 4: Technical verification

Before clinical use, the candidate protocol is run on a structured phantom to confirm geometric accuracy, the absence of new structured artefact and the behaviour of fine detail and low-contrast inserts at the chosen setting. Global similarity indices such as the structural similarity index are convenient for comparing reconstructions of the same data (Wang et al., 2004), but they summarise the whole image and are insensitive to the loss of a small lesion, so they are recorded without being treated as evidence of diagnostic adequacy. A small number of volunteer examinations, under the arrangements the department has for volunteer scanning, then establishes that the shortened sequence runs as intended and that the breath-hold has in fact become shorter.

3.5. Stage 5: Blinded reader comparison

The central test is a paired comparison in clinical examinations. Each person undergoes the reference sequence and the candidate sequence in the same session, in an order that is varied to avoid a systematic effect of fatigue or contrast timing. For post-contrast sequences, where both cannot occupy the same phase, the comparison is restricted to phases that can be duplicated, or to consecutive examinations of the same person. At least two radiologists, unaware of the method, score each series on predefined scales for motion artefact, sharpness, noise, artificial texture and overall diagnostic quality, and separately record the focal lesions they detect. Lesion detection against a reference reading matters more than preference scores, because readers tend to prefer smooth images. The sample size, the scales and the analysis are fixed in a written plan, and the work is reported in line with the CLAIM checklist (Mongan et al., 2020).

3.6. Stage 6: Monitoring in routine use

If the comparison supports adoption, the candidate protocol replaces the reference for the stated indication, and monitoring begins. Two simple measures are proposed. One is the proportion of examinations in which a sequence is repeated or flagged by the reporting radiologist as limited by motion, tracked before and after the change. The other is a log of discrepancies in which a finding on follow-up or on another modality was not visible on the deep learning images, each reviewed against the conventional reconstruction if it was kept. Monitoring is renewed after any software update or change of coil.

4. Fitting the Procedure to a Department

The procedure can be scaled. A department with limited research capacity may confine itself to one sequence, perform the technical checks, and run the reader comparison on a modest consecutive series as a service evaluation under local governance. A larger centre can add free-breathing radial acquisitions, quantitative sequences and formal lesion-detection analysis. In both, the essential discipline is the same: one change at a time, a retained reference, fixed settings and readers who do not know which image they are scoring. Radiographers need to be told what has changed and why, since the benefit depends on the shorter breath-hold being coached and used. Old instructions given for a new sequence gain nothing.

5. Limitations and What Remains to Be Shown

The procedure is a proposal and has not been tested. Whether its stages are sufficient, or more than a busy department will carry out, can be learned only by using it. Its reader comparison measures image quality and lesion detection within a single session and cannot show effects on diagnosis or clinical outcome. It also inherits the difficulty of any evaluation of learned reconstruction: the absence of a ground truth. In clinical examinations there is no motion-free, fully sampled image against which to judge, and agreement between two reconstructions of the same motion-affected data proves little.

Three further matters are open. The first is failure behaviour. When motion is severe, a network trained on motion-free data is operating outside its training distribution, and whether it then degrades visibly or produces a clean-looking but wrong image is the question that most needs study ([Antun et al., 2020](#)). The second is generalisation across field strengths, coils, contrast agents and body habitus, which product documentation seldom describes in detail. The third is the place of direct motion-correction networks. If these reach clinical products, the procedure would need an additional stage to test them on data with known, deliberately induced motion ([Haskell et al., 2019](#)). None of this argues against adoption. It argues for adoption with the reference images kept and the results watched.

6. Conclusion

Deep learning reconstruction reduces motion artefact in abdominal MRI mainly by making short acquisitions usable, and it does so only when protocols are redesigned to shorten the time for which stillness is required. The six-stage procedure proposed here gives a department an orderly way to make that change: define the problem, alter the acquisition one step at a time, fix the reconstruction settings, verify technically, compare under blinded reading with lesion detection as the main measure, and monitor afterwards. It treats the learned reconstruction as a component to be checked against the data, not as a guarantee of quality. The procedure has yet to be applied, and reports of its use, including instances where the new images proved worse, would be the most useful next contribution.

Conflict of interest

To be declared by the author(s) before publication.

Funding

To be declared by the author(s) before publication.

References

- Aggarwal, H.K., Mani, M.P. and Jacob, M. (2019). [MoDL: Model-based deep learning architecture for inverse problems](#). *IEEE Transactions on Medical Imaging*, 38(2): 394-405. <https://doi.org/10.1109/TMI.2018.2865356>
- Antun, V., Renna, F., Poon, C., Adcock, B. and Hansen, A.C. (2020). [On instabilities of deep learning in image reconstruction and the potential costs of AI](#). *Proceedings of the National Academy of Sciences*, 117(48): 30088-30095. <https://doi.org/10.1073/pnas.1907377117>
- Davenport, M.S., Viglianti, B.L., Al-Hawary, M.M., et al. (2013). [Comparison of acute transient dyspnea](#)

- after intravenous administration of gadoxetate disodium and gadobenate dimeglumine: Effect on arterial phase image quality. *Radiology*, 266(2): 452-461.
- Feng, L., Axel, L., Chandarana, H., Block, K.T., Sodickson, D.K. and Otazo, R. (2016). XD-GRASP: Golden-angle radial MRI with reconstruction of extra motion-state dimensions using compressed sensing. *Magnetic Resonance in Medicine*, 75(2): 775-788. <https://doi.org/10.1002/mrm.25665>
- Feng, L., Grimm, R., Block, K.T., et al. (2014). Golden-angle radial sparse parallel MRI: Combination of compressed sensing, parallel imaging, and golden-angle radial sampling for fast and flexible dynamic volumetric MRI. *Magnetic Resonance in Medicine*, 72(3): 707-717. <https://doi.org/10.1002/mrm.24980>
- Griswold, M.A., Jakob, P.M., Heidemann, R.M., et al. (2002). Generalized autocalibrating partially parallel acquisitions (GRAPPA). *Magnetic Resonance in Medicine*, 47(6): 1202-1210. <https://doi.org/10.1002/mrm.10171>
- Hammernik, K., Klatzer, T., Kobler, E., Recht, M.P., Sodickson, D.K., Pock, T. and Knoll, F. (2018). Learning a variational network for reconstruction of accelerated MRI data. *Magnetic Resonance in Medicine*, 79(6): 3055-3071. <https://doi.org/10.1002/mrm.26977>
- Haskell, M.W., Cauley, S.F., Bilgic, B., et al. (2019). Network Accelerated Motion Estimation and Reduction (NAMER): Convolutional neural network guided retrospective motion correction using a separable motion model. *Magnetic Resonance in Medicine*, 82(4): 1452-1461.
- Knoll, F., Hammernik, K., Zhang, C., et al. (2020). Deep-learning methods for parallel magnetic resonance imaging reconstruction: A survey of the current approaches, trends, and prospects. *IEEE Signal Processing Magazine*, 37(1): 128-140.
- Küstner, T., Armanious, K., Yang, J., Yang, B., Schick, F. and Gatidis, S. (2019). Retrospective correction of motion-affected MR images using deep learning frameworks. *Magnetic Resonance in Medicine*, 82(4): 1527-1540.
- Lustig, M., Donoho, D. and Pauly, J.M. (2007). Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magnetic Resonance in Medicine*, 58(6): 1182-1195. <https://doi.org/10.1002/mrm.21391>
- Mongan, J., Moy, L. and Kahn, C.E. (2020). Checklist for Artificial Intelligence in Medical Imaging (CLAIM): A guide for authors and reviewers. *Radiology: Artificial Intelligence*, 2(2): e200029. <https://doi.org/10.1148/ryai.2020200029>
- Pipe, J.G. (1999). Motion correction with PROPELLER MRI: Application to head motion and free-breathing cardiac imaging. *Magnetic Resonance in Medicine*, 42(5): 963-969.
- Pruessmann, K.P., Weiger, M., Scheidegger, M.B. and Boesiger, P. (1999). SENSE: Sensitivity encoding for fast MRI. *Magnetic Resonance in Medicine*, 42(5): 952-962.
- Schlemper, J., Caballero, J., Hajnal, J.V., Price, A.N. and Rueckert, D. (2018). A deep cascade of convolutional neural networks for dynamic MR image reconstruction. *IEEE Transactions on Medical Imaging*, 37(2): 491-503. <https://doi.org/10.1109/TMI.2017.2760978>
- Wang, Z., Bovik, A.C., Sheikh, H.R. and Simoncelli, E.P. (2004). Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4): 600-612. <https://doi.org/10.1109/TIP.2003.819861>
- Zaitsev, M., Maclaren, J. and Herbst, M. (2015). Motion artifacts in MRI: A complex problem with many partial solutions. *Journal of Magnetic Resonance Imaging*, 42(4): 887-901. <https://doi.org/10.1002/jmri.24850>
- Zbontar, J., Knoll, F., Sriram, A., et al. (2018). fastMRI: An open dataset and benchmarks for accelerated MRI. arXiv preprint arXiv:1811.08839.

Cite this article as: Olsen M., Farouk S. and Milosevic L. (2026). Motion Artefact Reduction in Abdominal Magnetic Resonance Imaging Using Deep Learning Reconstruction. *International journal of Radiology and Medical Imaging Technology*. 1(1), 21-26. doi: 10.62715/ijrmit.2026.v1.i1.pp21-26