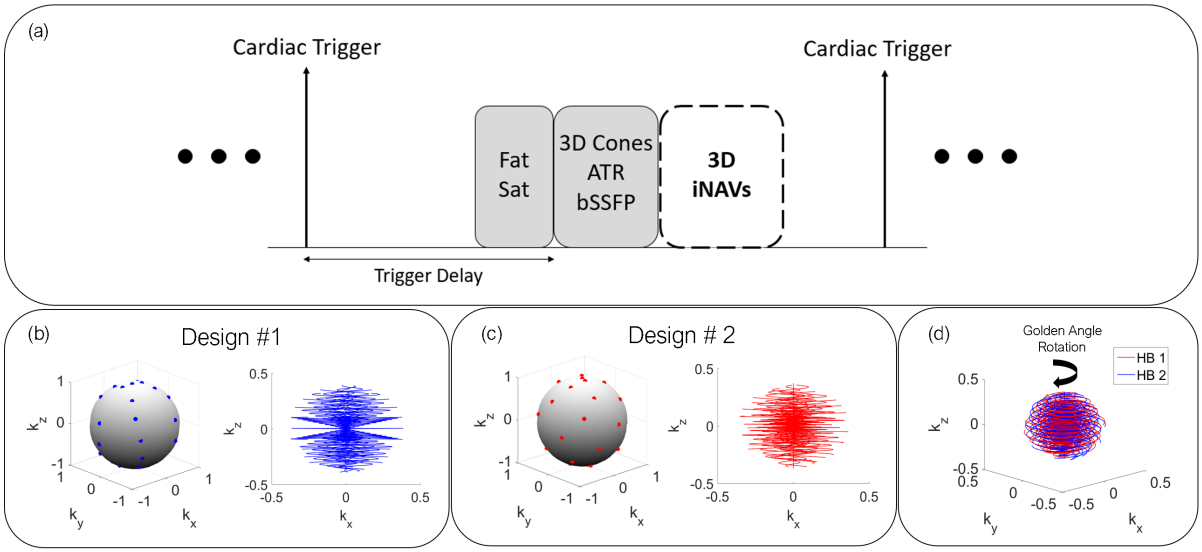
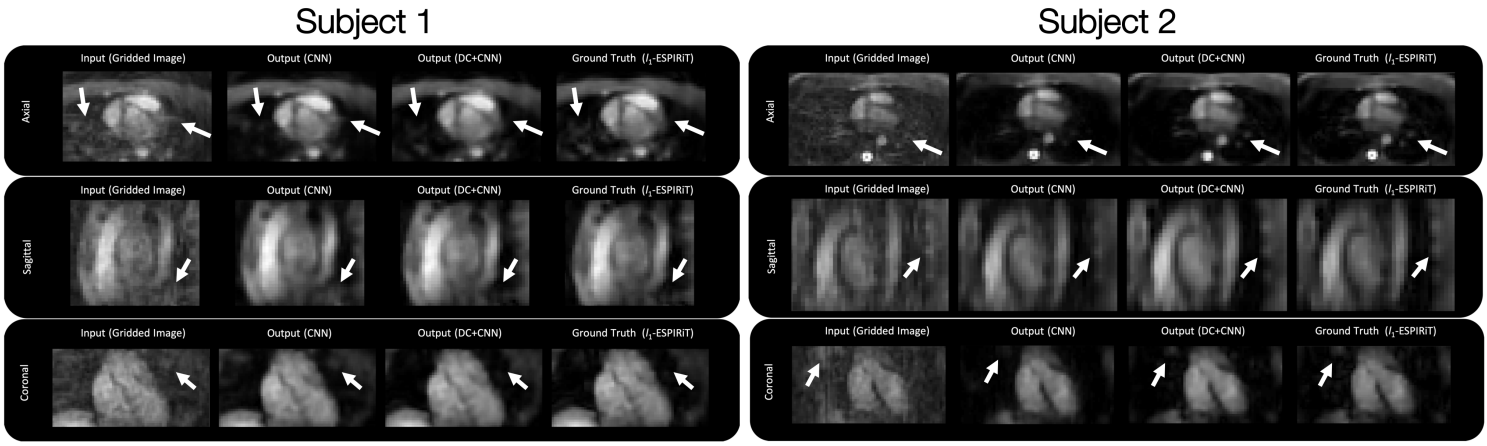
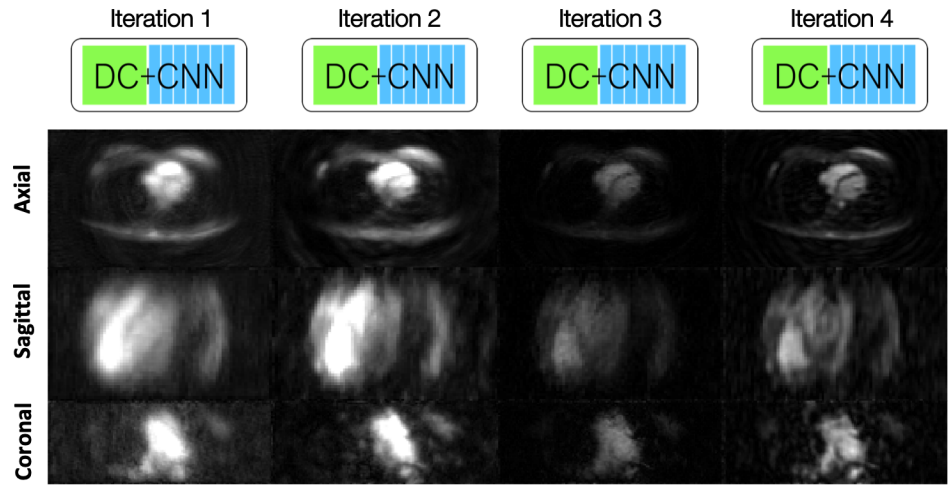


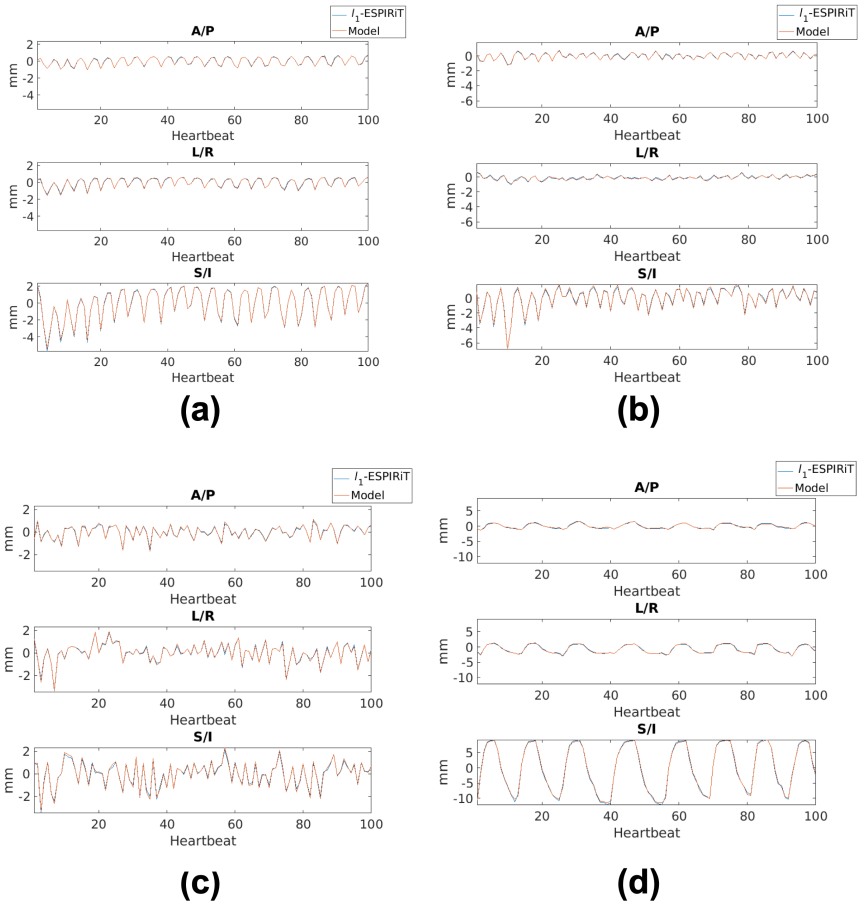
Supporting Information Figure S1: (a) For the free-breathing CMRA acquisition scheme, the 3D iNAVs are collected every heartbeat following the fat saturation and imaging data acquisition as shown in the timing diagram. The 3D iNAVs are acquired using a variable-density, undersampled 3D cones trajectory. (b) The first design uses a sequential-based acquisition with multiple readouts (and uniform azimuthal rotations) within each conical surface. (c) The second design employs a phyllotaxis scheme with unique conical surfaces and golden angle azimuthal rotations. The blue and red points on the unit sphere represent the polar angles for each corresponding cone readout. (d) In addition, some of the datasets rotate the two trajectory designs between heartbeats by the golden angle to help the model further generalize during training.



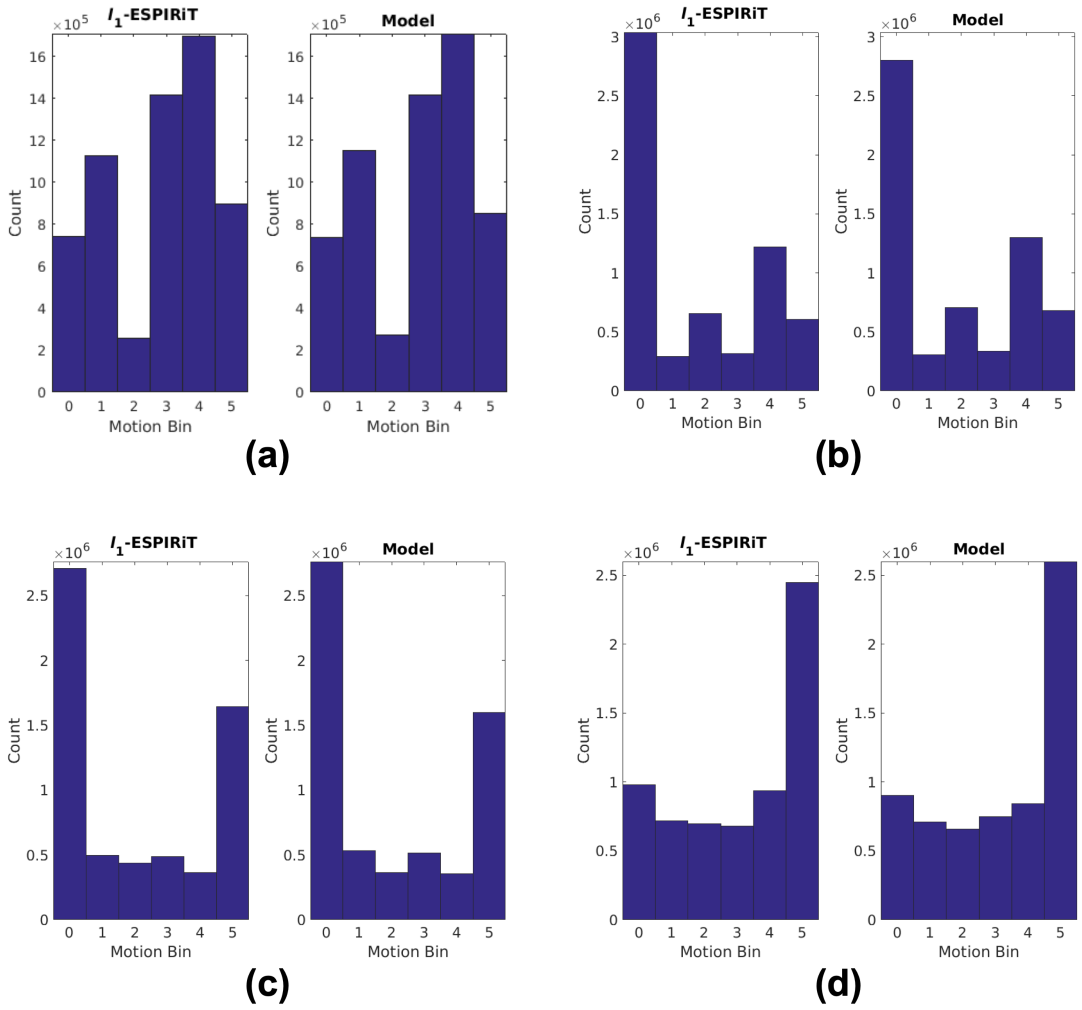
Supporting Information Figure S2: (The axial, sagittal, and coronal slices are shown from one heartbeat. 3D iNAV inputs (gridded images using the NUFFT operator), outputs using CNNs without data-consistency (DC), outputs using the unrolled model with DC and CNNs, and ground truths (l_1-ESPIRiT) are shown respectively for subjects 1-2 (test datasets). For each method, the differences are highlighted using white arrows which show how the unrolled model more closely matches the ground truth compared to using a CNN without DC.



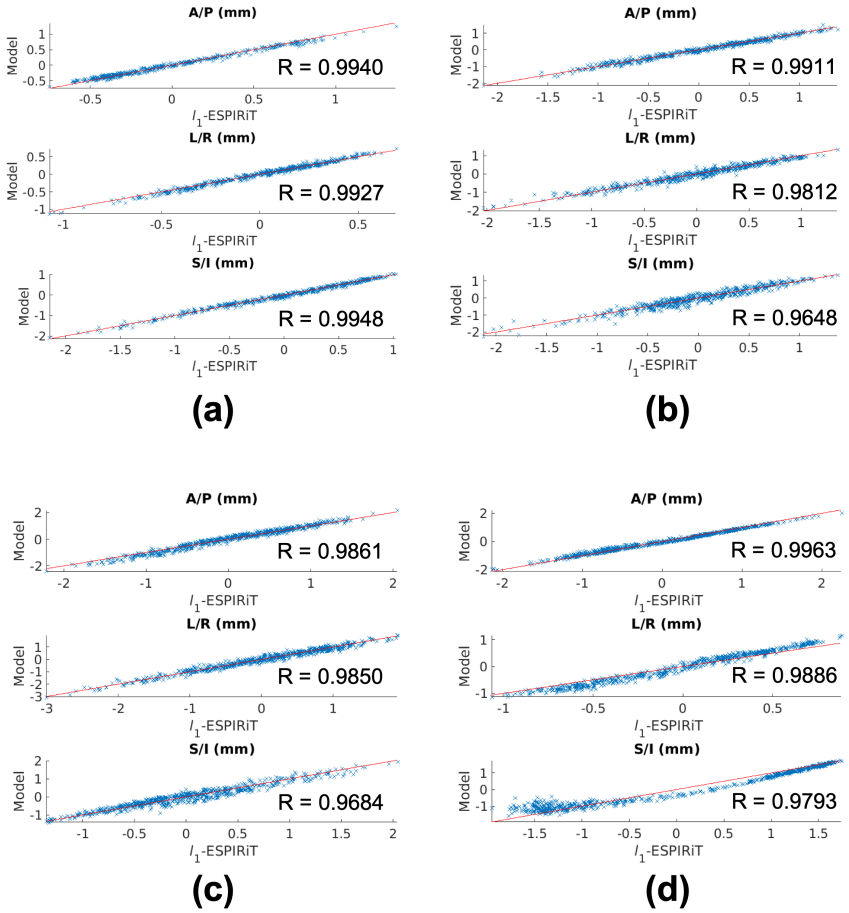
Supporting Information Figure S3: (a) The outputs of each iteration in the unrolled model during training are shown for one example dataset. The respective outputs for each of the 4 iterations (gradient steps) highlight the behavior of each different blocks in the model. For each iteration, image depiction is improved by enhancing the structures throughout the axial, sagittal, and coronal slices.



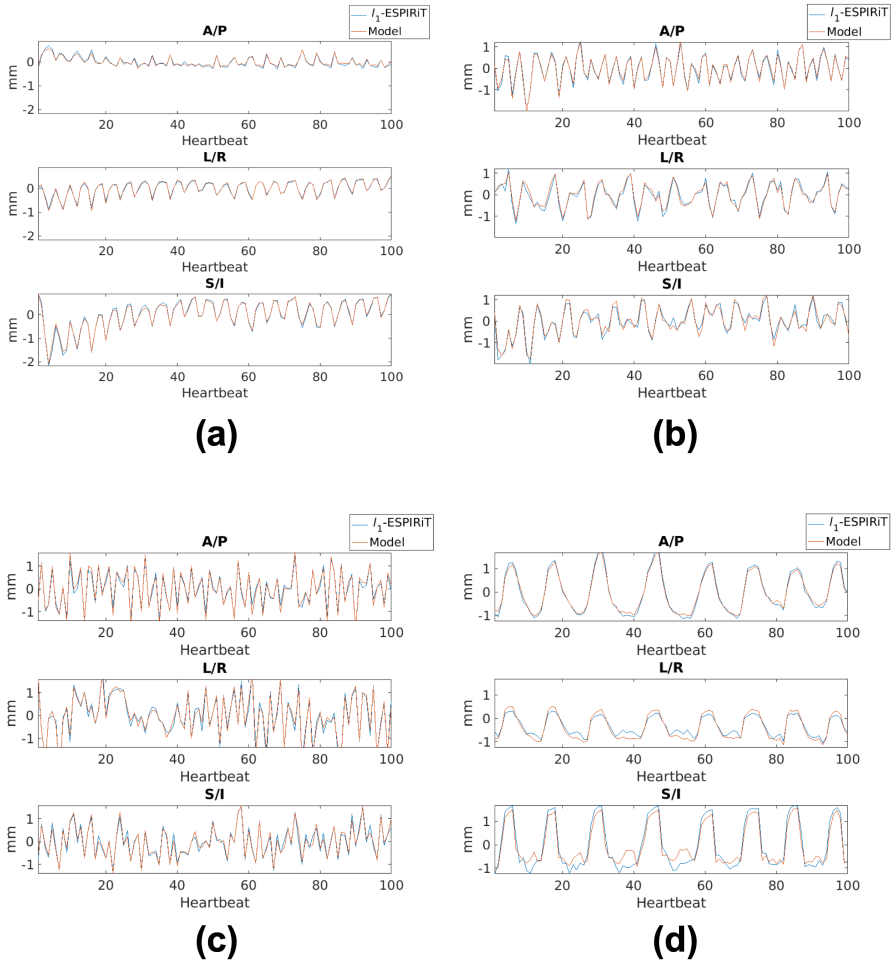
Supporting Information Figure S4: The global motion estimates (first 100 heartbeats) generated from l_1-ESPIRiT, and the DL model-based 3D iNAVs. The plots show how the motion estimates extracted from the l_1-ESPIRiT and DL model-based 3D iNAVs track similar motion in all directions for all four subjects (a-d).



Supporting Information Figure S5: The histograms generated from the outcomes of the autofocusing algorithm when using l_1-ESPIRiT, and the DL model-based 3D iNAVs for subjects 1-4 (a-d). The histograms show the global and residual motion bins (0-5), respectively. For subjects 2 and 3 (b,c), the global bin is the most selected by autofocusing which shows that there was less residual motion beyond the rigid-body translational motion. For subjects 1 and 4 (a,d), bins four and five are the most selected, demonstrating that there was additional residual motion beyond translational.



Supporting Information Figure S6: (a) Example residual motion estimate scatter plots and correlation coefficients (R) for all four subjects (a-d) generated from l_1-ESPIRiT, and the DL model-based 3D iNAVs. For subjects 1 and 2, bin 4 is shown, and for subjects 3 and 4, bin 5 is shown. The scatter plots show slightly less correlation compared to the global estimates (Figure 4) which may partly be attributed to minor interpolation differences between the l_1-ESPIRiT, and the DL model-based reconstructions.



Supporting Information Figure S7: (a) Example residual motion estimates (first 100 heartbeats) for all four subjects (a-d) generated from l_1-ESPIRiT and the DL model-based 3D iNAVs. The corresponding motion bin estimates from Supporting Information Figure S5 are shown. The residual motion estimates (A/P, L/R, S/I) allow for residual motion correction which the global translations do not fully capture.

l_1-ESPIRiT & Model Correlations	Direction	Motion Bin					
		Global	1	2	3	4	5
Subject 1	A/P	0.9961	0.9872	0.9966	0.9994	0.9940	0.9873
	L/R	0.9980	0.9983	0.9931	0.9945	0.9927	0.9925
	S/I	0.9996	0.9997	0.9991	0.9988	0.9948	0.9997
Subject 2	A/P	0.9951	0.9732	0.9895	0.9764	0.9911	0.9672
	L/R	0.9914	0.9794	0.9545	0.9834	0.9812	0.9720
	S/I	0.9966	0.9451	0.9747	0.9951	0.9648	0.9481
Subject 3	A/P	0.9958	0.9793	0.9926	0.9930	0.9909	0.9861
	L/R	0.9962	0.9499	0.9907	0.9829	0.9094	0.9850
	S/I	0.9961	0.9418	0.9800	0.9762	0.9108	0.9684
Subject 4	A/P	0.9929	0.9978	0.9551	0.9444	0.9687	0.9963
	L/R	0.9982	0.9977	0.9984	0.9886	0.9979	0.9886
	S/I	0.9992	0.9868	0.9983	0.9928	0.9332	0.9793

Supporting Information Table S1: The correlation coefficients between motion estimates (in A/P, L/R, and S/I) obtained from l_1-ESPIRiT and the DL model-based 3D iNAVs for the global and five spatial bins. These motion estimates are used to generate a bank of six 3D motion-compensated reconstructions (from one global motion estimate, and five residual localized motion estimates) used as candidates for the autofocusing algorithm.

Vessel Sharpness Comparison			
Subject	**Coronary**	**Method**	**IEPA (a.u.)**
Subject 1	RCA	l_1-ESPIRiT	0.1485
		Model	0.1481
	LAD	l_1-ESPIRiT	0.1288
		Model	0.1284
Subject 2	RCA	l_1-ESPIRiT	0.1361
		Model	0.1379
	LAD	l_1-ESPIRiT	0.1195
		Model	0.1204
Subject 3	RCA	l_1-ESPIRiT	0.1314
		Model	0.1311
	LAD	l_1-ESPIRiT	0.1082
		Model	0.1089
Subject 4	RCA	l_1-ESPIRiT	0.1244
		Model	0.1273
	LAD	l_1-ESPIRiT	0.0969
		Model	0.0982

Supporting Information Table S2: The average vessel sharpness values (IEPA) for all four subjects. The IEPA values were calculated on the motion-corrected images using the l_1-ESPIRiT and the DL model-based 3D iNAVs and along the RCA and LCA (50 mm segments). Similar IEPA values are shown using both methods which correspond to the similar image quality for the motion-corrected images in Figures 5 and 6.