

Additional file 1: Supplementary Material

Four-grade lumbar foraminal stenosis grading on sagittal T2-weighted MRI: a patient-disjoint comparison of ordinal models, auxiliary supervision, and fixed fusion

Supplementary contents

- Supplementary Methods S1. Cohort, fold identity, and leakage control
- Supplementary Methods S2. Frozen preprocessing, checkpoint, and fusion rules
- Supplementary Methods S3. Model and training implementation details
- Supplementary Results S1. Complete fold-wise performance
- Supplementary Results S2. Per-grade behaviour and pooled confusion detail
- Supplementary Results S3. Complete 20,000-replicate paired bootstrap results
- Supplementary Results S4. Modern matched-arm MSaux context
- Supplementary Interpretability S1. Full 20-case locked audit
- Supplementary Interpretability S2. Deletion faithfulness
- Supplementary Reproducibility S1. Frozen-source provenance and exact checkpoint identities

Supplementary Methods S1. Cohort, fold identity, and leakage control

The frozen cohort comprised 468 patients and 2,978 expert-defined sagittal T2-weighted foraminal ROIs: 2,009 Normal, 507 Mild, 254 Moderate, and 208 Severe. The final analysis used five patient-disjoint outer folds. Outer-fold identity was verified directly across all 14 reader-facing OOF sources using the common `protocol_v4_row_id` key: each OOF contained the same 2,978 keys, the direct `outer_fold` assignments agreed with the authoritative reference map for 2,978/2,978 ROIs, and there were zero cross-system fold conflicts. The raw cohort manifest was joined to this verified map by `roi_path` and remained patient-disjoint. The fixed inner validation mapping was 0->1, 1->2, 2->3, 3->4, and 4->0. Pooled OOF results concatenate the five held-out folds rather than averaging fold metrics.

Supplementary Table S1. Patient, ROI, and grade counts by verified outer fold.

Fold	Patients	ROIs	Normal	Mild	Moderate	Severe
0	94	618	418	106	52	42
1	94	596	402	102	49	43
2	94	557	372	95	50	40
3	93	616	427	96	52	41

Fold	Patients	ROIs	Normal	Mild	Moderate	Severe
4	93	591	390	108	51	42
Total	468	2978	2009	507	254	208

Fold identity source: consensus of all 14 reader-facing OOF files; reference key protocol_v4_row_id; raw-manifest join key roi_path; zero patient leakage.

Supplementary Methods S2. Frozen preprocessing, checkpoint, and fusion rules

The historical compact branch used its matched 64x64 single-channel grayscale pipeline. The frozen DeiT branch used 224x224 input with repeated grayscale channels. The modern ConvNeXt arm used its matched 224x224 grayscale-to-RGB/ImageNet-normalized preprocessing. Cross-family comparisons are therefore interpreted at the system level rather than as pure input-controlled effects. The heterogeneous fusion rule was prespecified as $P_{\text{final}} = 0.40 P_{\text{branch}} + 0.60 P_{\text{DeiT}}$, with no fusion-weight search. Checkpoint selection followed each model's frozen native objective without outer-fold tuning.

Input/preprocessing fairness and checkpoint rules were taken from the frozen Step06A4E evidence record; no preprocessing or checkpoint rule was changed for this Supplement.

Supplementary Methods S3. Model and training implementation details

The final roster contains 14 systems spanning source-canonical controls, compact nominal/ordinal/MSaux variants, fixed-DeiT fusions, and a representative modern ConvNeXt arm. MSaux denotes the training-only moderate-to-severe auxiliary objective $P(y \geq 2)$, with fixed weight 0.50 for the designated MSaux systems and no additional inference branch. This Supplement introduces no new model, threshold, checkpoint-selection rule, fusion weight, or training run.

System	Frozen OOF source identifier
CNN + fixed DeiT	Original_CNN_DeiT_60_40_native_loss_fixed_inner_OOF.csv
CORAL + fixed DeiT	CORAL_DeiT_60_40_native_loss_fixed_inner_OOF.csv
CNN-MSaux + fixed DeiT	CNN_Risk_DeiT_60_40_native_loss_fixed_inner_OOF.csv
CORAL-MSaux + fixed DeiT	CORAL_RISK_DeiT_60_40_native_loss_fixed_inner_OOF.csv
ConvNeXt-CORAL + fixed DeiT	convnext_coral_plus_fixed60_deit_OOF_predictions.csv
ConvNeXt-CORAL-MSaux + fixed DeiT	convnext_coral_msaux_plus_fixed60_deit_OOF_predictions.csv
Original DeiT	Original_DeiT_native_loss_fixed_inner_OOF.csv
Original CNN	Original_CNN_native_loss_fixed_inner_OOF.csv
CORAL	CORAL_native_loss_fixed_inner_OOF.csv
CNN-MSaux	CNN_Risk_native_loss_fixed_inner_OOF.csv
CORAL-MSaux	CORAL_RISK_native_loss_fixed_inner_OOF.csv
ConvNeXt-CE	ConvNeXt-CE_exact_OOF.csv
ConvNeXt-CORAL	convnext_coral_OOF_predictions.csv
ConvNeXt-CORAL-MSaux	convnext_coral_msaux_OOF_predictions.csv

All 14 reader-facing OOF sources were independently verified against the common protocol_v4_row_id reference map for 2,978/2,978 ROIs.

Supplementary Results S1. Complete fold-wise performance

Supplementary Table S2 reports all six primary metrics for each held-out outer fold and each final system. Before this table was assembled, the pooled six-metric values were recomputed from every frozen OOF file and matched the frozen main-manuscript Table 2 for all 14 systems after four-decimal rounding.

System	Fold	N	Accuracy	Balanced Accuracy	Macro-F1	Weighted-F1	QWK	MAE
CNN + fixed DeiT	0	618	0.7120	0.6041	0.5799	0.7272	0.7853	0.3155
CNN + fixed DeiT	1	596	0.7466	0.6957	0.6673	0.7656	0.8162	0.2735
CNN + fixed DeiT	2	557	0.7289	0.6449	0.6257	0.7480	0.7884	0.2998
CNN + fixed DeiT	3	616	0.7110	0.6035	0.5894	0.7315	0.7693	0.3182
CNN + fixed DeiT	4	591	0.6870	0.6267	0.5969	0.7088	0.7619	0.3469
CORAL + fixed DeiT	0	618	0.7330	0.6078	0.5872	0.7397	0.7891	0.2994
CORAL + fixed DeiT	1	596	0.7685	0.6684	0.6511	0.7764	0.8273	0.2534
CORAL + fixed DeiT	2	557	0.7469	0.6253	0.6164	0.7555	0.7949	0.2837
CORAL + fixed DeiT	3	616	0.7256	0.5835	0.5719	0.7346	0.7794	0.3052
CORAL + fixed DeiT	4	591	0.7242	0.6183	0.5999	0.7260	0.7651	0.3215
CNN-MSaux + fixed DeiT	0	618	0.7055	0.6028	0.5791	0.7225	0.7799	0.3220
CNN-MSaux + fixed DeiT	1	596	0.7550	0.6858	0.6534	0.7685	0.8089	0.2735
CNN-MSaux + fixed DeiT	2	557	0.7163	0.6343	0.6110	0.7372	0.7844	0.3124
CNN-MSaux + fixed DeiT	3	616	0.7127	0.5901	0.5778	0.7318	0.7699	0.3166
CNN-MSaux + fixed DeiT	4	591	0.6853	0.6346	0.6003	0.7083	0.7625	0.3486
CORAL-MSaux + fixed DeiT	0	618	0.7460	0.6256	0.6051	0.7533	0.7954	0.2864
CORAL-MSaux + fixed DeiT	1	596	0.7869	0.6680	0.6554	0.7809	0.8245	0.2450
CORAL-MSaux + fixed DeiT	2	557	0.7487	0.6299	0.6234	0.7589	0.8051	0.2765
CORAL-MSaux + fixed DeiT	3	616	0.7386	0.5939	0.5765	0.7393	0.7869	0.2938

Additional file 1 - Supplementary Material

System	Fold	N	Accuracy	Balanced Accuracy	Macro-F1	Weighted-F1	QWK	MAE
CORAL-MSaux + fixed DeiT	4	591	0.7259	0.6325	0.6097	0.7311	0.7752	0.3147
ConvNeXt-CORAL + fixed DeiT	0	618	0.7411	0.5996	0.5814	0.7381	0.7953	0.2913
ConvNeXt-CORAL + fixed DeiT	1	596	0.7819	0.6640	0.6448	0.7739	0.8206	0.2517
ConvNeXt-CORAL + fixed DeiT	2	557	0.7469	0.6128	0.5992	0.7480	0.7891	0.2908
ConvNeXt-CORAL + fixed DeiT	3	616	0.7532	0.6068	0.5842	0.7475	0.7932	0.2841
ConvNeXt-CORAL + fixed DeiT	4	591	0.7174	0.6014	0.5780	0.7208	0.7703	0.3266
ConvNeXt-CORAL-MSaux + fixed DeiT	0	618	0.7476	0.5956	0.5769	0.7397	0.7929	0.2880
ConvNeXt-CORAL-MSaux + fixed DeiT	1	596	0.7752	0.6624	0.6397	0.7682	0.8150	0.2601
ConvNeXt-CORAL-MSaux + fixed DeiT	2	557	0.7469	0.6204	0.6051	0.7492	0.7919	0.2908
ConvNeXt-CORAL-MSaux + fixed DeiT	3	616	0.7565	0.6157	0.5928	0.7497	0.7923	0.2825
ConvNeXt-CORAL-MSaux + fixed DeiT	4	591	0.7208	0.6010	0.5750	0.7244	0.7694	0.3266
Original DeiT	0	618	0.6990	0.6064	0.5738	0.7205	0.7672	0.3366
Original DeiT	1	596	0.7215	0.6809	0.6393	0.7450	0.7899	0.3070
Original DeiT	2	557	0.7002	0.6321	0.6036	0.7260	0.7609	0.3375
Original DeiT	3	616	0.6948	0.5986	0.5799	0.7195	0.7504	0.3409
Original DeiT	4	591	0.6616	0.6147	0.5748	0.6879	0.7349	0.3841
Original CNN	0	618	0.7492	0.5405	0.5154	0.7131	0.7864	0.2929
Original CNN	1	596	0.7534	0.4982	0.5180	0.6935	0.7339	0.3003
Original CNN	2	557	0.7253	0.4856	0.4983	0.6896	0.7222	0.3285
Original CNN	3	616	0.7581	0.5175	0.5060	0.7122	0.7601	0.2922
Original CNN	4	591	0.7394	0.5414	0.5541	0.7003	0.7881	0.2893
CORAL	0	618	0.7589	0.5563	0.5549	0.7411	0.7928	0.2767
CORAL	1	596	0.7819	0.6138	0.6281	0.7625	0.8197	0.2450

Additional file 1 - Supplementary Material

System	Fold	N	Accuracy	Balanced Accuracy	Macro-F1	Weighted-F1	QWK	MAE
CORAL	2	557	0.7325	0.5142	0.5284	0.7009	0.7391	0.3178
CORAL	3	616	0.7695	0.5674	0.5613	0.7441	0.7944	0.2662
CORAL	4	591	0.7293	0.5606	0.5577	0.7021	0.7801	0.3046
CNN-MSaux	0	618	0.7476	0.5323	0.5152	0.7092	0.7822	0.2929
CNN-MSaux	1	596	0.7500	0.5313	0.5085	0.7135	0.7972	0.2886
CNN-MSaux	2	557	0.7217	0.4991	0.5015	0.6924	0.7161	0.3429
CNN-MSaux	3	616	0.7614	0.5652	0.5442	0.7307	0.7714	0.2873
CNN-MSaux	4	591	0.7225	0.5123	0.5293	0.6895	0.7496	0.3164
CORAL-MSaux	0	618	0.7557	0.5681	0.5648	0.7382	0.8112	0.2686
CORAL-MSaux	1	596	0.7651	0.5519	0.5476	0.7118	0.7903	0.2752
CORAL-MSaux	2	557	0.7469	0.5140	0.5484	0.7291	0.7398	0.2962
CORAL-MSaux	3	616	0.7679	0.5712	0.5383	0.7358	0.7803	0.2841
CORAL-MSaux	4	591	0.7563	0.6114	0.6085	0.7361	0.7971	0.2792
ConvNeXt-CE	0	618	0.7168	0.5789	0.5652	0.7317	0.7713	0.3139
ConvNeXt-CE	1	596	0.7903	0.6422	0.6569	0.7928	0.8141	0.2383
ConvNeXt-CE	2	557	0.7074	0.5900	0.5915	0.7227	0.7641	0.3232
ConvNeXt-CE	3	616	0.7614	0.5989	0.5838	0.7559	0.7808	0.2841
ConvNeXt-CE	4	591	0.7107	0.6181	0.6023	0.7248	0.7497	0.3299
ConvNeXt-CORAL	0	618	0.7411	0.5019	0.4361	0.6637	0.7393	0.3269
ConvNeXt-CORAL	1	596	0.7466	0.5068	0.4674	0.6690	0.7116	0.3339
ConvNeXt-CORAL	2	557	0.7217	0.4710	0.3697	0.6177	0.6827	0.3788
ConvNeXt-CORAL	3	616	0.7484	0.5018	0.4199	0.6643	0.7055	0.3425
ConvNeXt-CORAL	4	591	0.7191	0.4970	0.4524	0.6465	0.6926	0.3655
ConvNeXt-CORAL-MSaux	0	618	0.7282	0.4672	0.3593	0.6283	0.7015	0.3641
ConvNeXt-CORAL-MSaux	1	596	0.7315	0.4684	0.3688	0.6316	0.6967	0.3658
ConvNeXt-CORAL-MSaux	2	557	0.7217	0.4710	0.3636	0.6196	0.6923	0.3788
ConvNeXt-CORAL-MSaux	3	616	0.7468	0.4843	0.3735	0.6508	0.6778	0.3653
ConvNeXt-CORAL-MSaux	4	591	0.7073	0.4804	0.3546	0.6067	0.6915	0.3976

Scientific gate: pooled Table 2 parity 14/14 PASS; direct-fold versus authoritative-reference agreement 14 x 2,978/2,978 PASS.

Supplementary Results S2. Per-grade behaviour and complete pooled confusion detail

Supplementary Table S3. Per-grade recall from the frozen pooled OOF predictions.

System	Normal Recall	Mild Recall	Moderate Recall	Severe Recall
CNN + fixed DeiT	0.7934	0.5128	0.5276	0.7067
CORAL + fixed DeiT	0.8542	0.4162	0.4961	0.7163
CNN-MSaux + fixed DeiT	0.7939	0.5069	0.5079	0.7115
CORAL-MSaux + fixed DeiT	0.8676	0.4043	0.5000	0.7500
ConvNeXt-CORAL + fixed DeiT	0.8830	0.3432	0.4646	0.7788
ConvNeXt-CORAL-MSaux + fixed DeiT	0.8860	0.3314	0.4764	0.7837
Original DeiT	0.7631	0.5030	0.5394	0.7019
Original CNN	0.9642	0.1440	0.2283	0.7308
CORAL	0.9383	0.2485	0.3465	0.7163
CNN-MSaux	0.9438	0.1874	0.2559	0.7260
CORAL-MSaux	0.9462	0.2347	0.3425	0.7308
ConvNeXt-CE	0.8457	0.5030	0.4173	0.6587
ConvNeXt-CORAL	0.9806	0.0256	0.1063	0.8702
ConvNeXt-CORAL-MSaux	0.9836	0.0000	0.0000	0.9135

Supplementary Table S4. Complete pooled OOF confusion counts for all 14 systems.

System	True Grade	Pred Normal	Pred Mild	Pred Moderate	Pred Severe	Support
CNN + fixed DeiT	Normal	1594	363	49	3	2009
CNN + fixed DeiT	Mild	131	260	103	13	507
CNN + fixed DeiT	Moderate	9	72	134	39	254
CNN + fixed DeiT	Severe	0	6	55	147	208
CORAL + fixed DeiT	Normal	1716	240	50	3	2009
CORAL + fixed DeiT	Mild	177	211	105	14	507
CORAL + fixed DeiT	Moderate	17	65	126	46	254
CORAL + fixed DeiT	Severe	3	3	53	149	208
CNN-MSaux + fixed DeiT	Normal	1595	360	50	4	2009
CNN-MSaux + fixed DeiT	Mild	130	257	107	13	507

Additional file 1 - Supplementary Material

System	True Grade	Pred Normal	Pred Mild	Pred Moderate	Pred Severe	Support
CNN-MSaux + fixed DeiT	Moderate	11	69	129	45	254
CNN-MSaux + fixed DeiT	Severe	1	4	55	148	208
CORAL-MSaux + fixed DeiT	Normal	1743	215	47	4	2009
CORAL-MSaux + fixed DeiT	Mild	185	205	101	16	507
CORAL-MSaux + fixed DeiT	Moderate	18	64	127	45	254
CORAL-MSaux + fixed DeiT	Severe	2	4	46	156	208
ConvNeXt-CORAL + fixed DeiT	Normal	1774	177	54	4	2009
ConvNeXt-CORAL + fixed DeiT	Mild	211	174	102	20	507
ConvNeXt-CORAL + fixed DeiT	Moderate	19	59	118	58	254
ConvNeXt-CORAL + fixed DeiT	Severe	3	3	40	162	208
ConvNeXt-CORAL-MSaux + fixed DeiT	Normal	1780	167	57	5	2009
ConvNeXt-CORAL-MSaux + fixed DeiT	Mild	210	168	109	20	507
ConvNeXt-CORAL-MSaux + fixed DeiT	Moderate	21	55	121	57	254
ConvNeXt-CORAL-MSaux + fixed DeiT	Severe	2	4	39	163	208
Original DeiT	Normal	1533	403	68	5	2009
Original DeiT	Mild	116	255	120	16	507
Original DeiT	Moderate	8	66	137	43	254
Original DeiT	Severe	0	7	55	146	208
Original CNN	Normal	1937	58	12	2	2009
Original CNN	Mild	372	73	47	15	507
Original CNN	Moderate	75	65	58	56	254
Original CNN	Severe	7	16	33	152	208
CORAL	Normal	1885	103	19	2	2009

System	True Grade	Pred Normal	Pred Mild	Pred Moderate	Pred Severe	Support
CORAL	Mild	305	126	64	12	507
CORAL	Moderate	52	60	88	54	254
CORAL	Severe	4	13	42	149	208
CNN-MSaux	Normal	1896	90	18	5	2009
CNN-MSaux	Mild	328	95	67	17	507
CNN-MSaux	Moderate	69	45	65	75	254
CNN-MSaux	Severe	7	9	41	151	208
CORAL-MSaux	Normal	1901	82	23	3	2009
CORAL-MSaux	Mild	308	119	65	15	507
CORAL-MSaux	Moderate	52	60	87	55	254
CORAL-MSaux	Severe	4	12	40	152	208
ConvNeXt-CE	Normal	1699	265	40	5	2009
ConvNeXt-CE	Mild	168	255	68	16	507
ConvNeXt-CE	Moderate	23	84	106	41	254
ConvNeXt-CE	Severe	2	12	57	137	208
ConvNeXt-CORAL	Normal	1970	6	10	23	2009
ConvNeXt-CORAL	Mild	425	13	21	48	507
ConvNeXt-CORAL	Moderate	113	12	27	102	254
ConvNeXt-CORAL	Severe	16	3	8	181	208
ConvNeXt-CORAL-MSaux	Normal	1976	0	0	33	2009
ConvNeXt-CORAL-MSaux	Mild	423	0	2	82	507
ConvNeXt-CORAL-MSaux	Moderate	118	0	0	136	254
ConvNeXt-CORAL-MSaux	Severe	18	0	0	190	208

Supplementary Results S3. Complete 20,000-replicate paired bootstrap results

Supplementary Table S5 reports the complete frozen patient-cluster paired-bootstrap comparison of CORAL-MSaux + fixed DeiT against the five fixed-fusion comparators. The resampling unit was the patient (20,000 replicates; seed 42). Effects use the frozen source orientation: positive values favor CORAL-MSaux + fixed DeiT, including the direction-oriented MAE effect. No bootstrap was rerun during Supplement assembly.

Comparator	Metric	Oriented delta	95% CI	CI excludes 0
CNN + fixed DeiT	Accuracy	0.0322	[0.0207, 0.0438]	Yes
CNN + fixed DeiT	Balanced accuracy	-0.0047	[-0.0186, 0.0090]	No

Additional file 1 - Supplementary Material

Comparator	Metric	Oriented delta	95% CI	CI excludes 0
CNN + fixed DeiT	Macro-F1	0.0031	[-0.0100, 0.0159]	No
CNN + fixed DeiT	Weighted F1	0.0166	[0.0059, 0.0271]	Yes
CNN + fixed DeiT	QWK	0.0132	[0.0045, 0.0219]	Yes
CNN + fixed DeiT	MAE	0.0275	[0.0154, 0.0397]	Yes
CORAL + fixed DeiT	Accuracy	0.0097	[0.0024, 0.0172]	Yes
CORAL + fixed DeiT	Balanced accuracy	0.0098	[-0.0006, 0.0201]	No
CORAL + fixed DeiT	Macro-F1	0.0094	[-0.0006, 0.0193]	No
CORAL + fixed DeiT	Weighted F1	0.0065	[-0.0006, 0.0136]	No
CORAL + fixed DeiT	QWK	0.0063	[-0.0000, 0.0129]	No
CORAL + fixed DeiT	MAE	0.0094	[0.0014, 0.0175]	Yes
CNN-MSaux + fixed DeiT	Accuracy	0.0343	[0.0234, 0.0453]	Yes
CNN-MSaux + fixed DeiT	Balanced accuracy	0.0004	[-0.0126, 0.0135]	No
CNN-MSaux + fixed DeiT	Macro-F1	0.0098	[-0.0023, 0.0219]	No
CNN-MSaux + fixed DeiT	Weighted F1	0.0189	[0.0088, 0.0290]	Yes
CNN-MSaux + fixed DeiT	QWK	0.0161	[0.0074, 0.0248]	Yes
CNN-MSaux + fixed DeiT	MAE	0.0312	[0.0198, 0.0428]	Yes
ConvNeXt-CORAL + fixed DeiT	Accuracy	0.0010	[-0.0084, 0.0102]	No
ConvNeXt-CORAL + fixed DeiT	Balanced accuracy	0.0131	[-0.0007, 0.0267]	No
ConvNeXt-CORAL + fixed DeiT	Macro-F1	0.0170	[0.0038, 0.0303]	Yes
ConvNeXt-CORAL + fixed DeiT	Weighted F1	0.0070	[-0.0020, 0.0161]	No
ConvNeXt-CORAL + fixed DeiT	QWK	0.0035	[-0.0046, 0.0120]	No
ConvNeXt-CORAL + fixed DeiT	MAE	0.0054	[-0.0048, 0.0155]	No
ConvNeXt-CORAL-MSaux + fixed DeiT	Accuracy	-0.0003	[-0.0100, 0.0092]	No
ConvNeXt-CORAL-MSaux + fixed DeiT	Balanced accuracy	0.0111	[-0.0036, 0.0259]	No
ConvNeXt-CORAL-MSaux + fixed DeiT	Macro-F1	0.0168	[0.0028, 0.0309]	Yes
ConvNeXt-CORAL-MSaux + fixed DeiT	Weighted F1	0.0065	[-0.0030, 0.0159]	No
ConvNeXt-CORAL-MSaux + fixed DeiT	QWK	0.0050	[-0.0030, 0.0135]	No
ConvNeXt-CORAL-MSaux + fixed DeiT	MAE	0.0060	[-0.0041, 0.0165]	No

Supplementary Results S4. Modern matched-arm MSaux context

Within the modern matched ConvNeXt arm, the MSaux effect was system-context dependent. Standalone ConvNeXt-CORAL showed an adverse MSaux effect on the frozen pooled OOF metrics. With the same fixed DeiT fusion rule, that adverse standalone effect was attenuated. This is reported as context dependence and attenuation, not as a universal MSaux benefit or fusion synergy.

System	Accuracy	Balanced Accuracy	Macro-F1	Weighted-F1	QWK	MAE
ConvNeXt-CORAL	0.7357	0.4957	0.4325	0.6539	0.7067	0.3489
ConvNeXt-CORAL-MSaux	0.7273	0.4743	0.3638	0.6276	0.6921	0.3741
ConvNeXt-CORAL + fixed DeiT	0.7482	0.6174	0.5985	0.7459	0.7939	0.2888
ConvNeXt-CORAL-MSaux + fixed DeiT	0.7495	0.6194	0.5987	0.7465	0.7924	0.2895

Supplementary Interpretability S1. Full 20-case locked audit

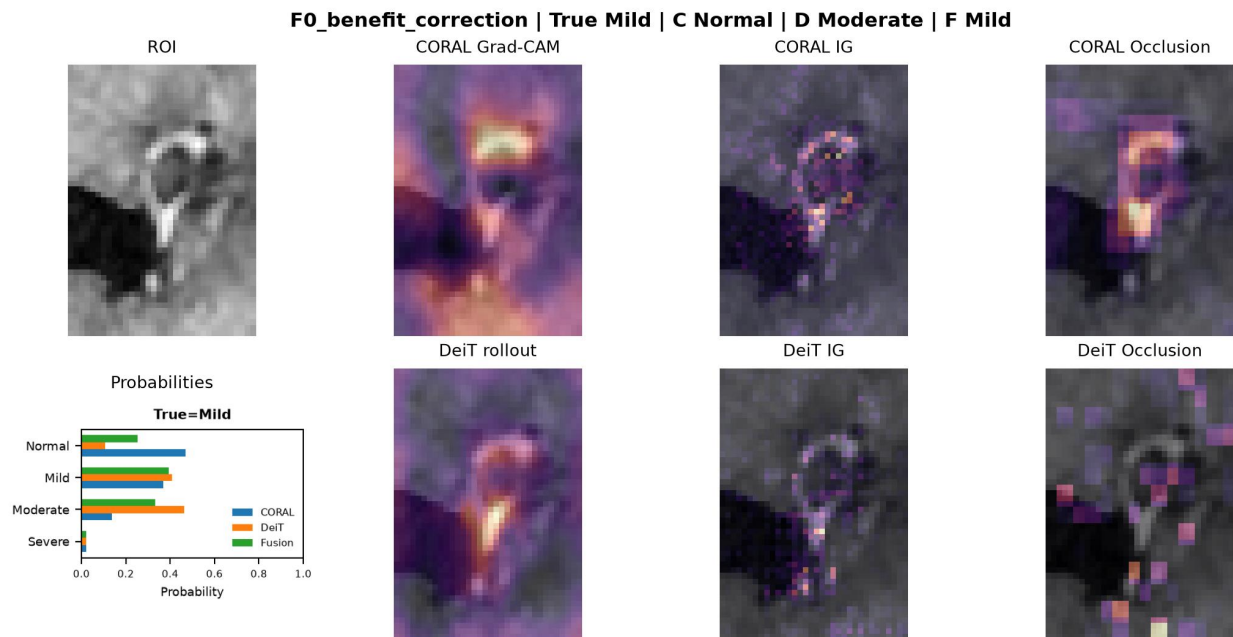
The interpretability audit used 20 unique patients, comprising four prespecified roles per outer fold. The ten fold-specific checkpoints were reverified by SHA, and prediction parity across CORAL-MSaux, Original DeiT, and fixed fusion was 60/60. CORAL-MSaux used Grad-CAM, integrated gradients, and occlusion; Original DeiT used attention rollout, integrated gradients, and occlusion. No synthetic fused saliency was created. Four of the 20 CORAL Grad-CAM maps were near-zero and are retained without enhancement or post hoc case replacement.

Supplementary Table S7. Locked 20-case metadata.

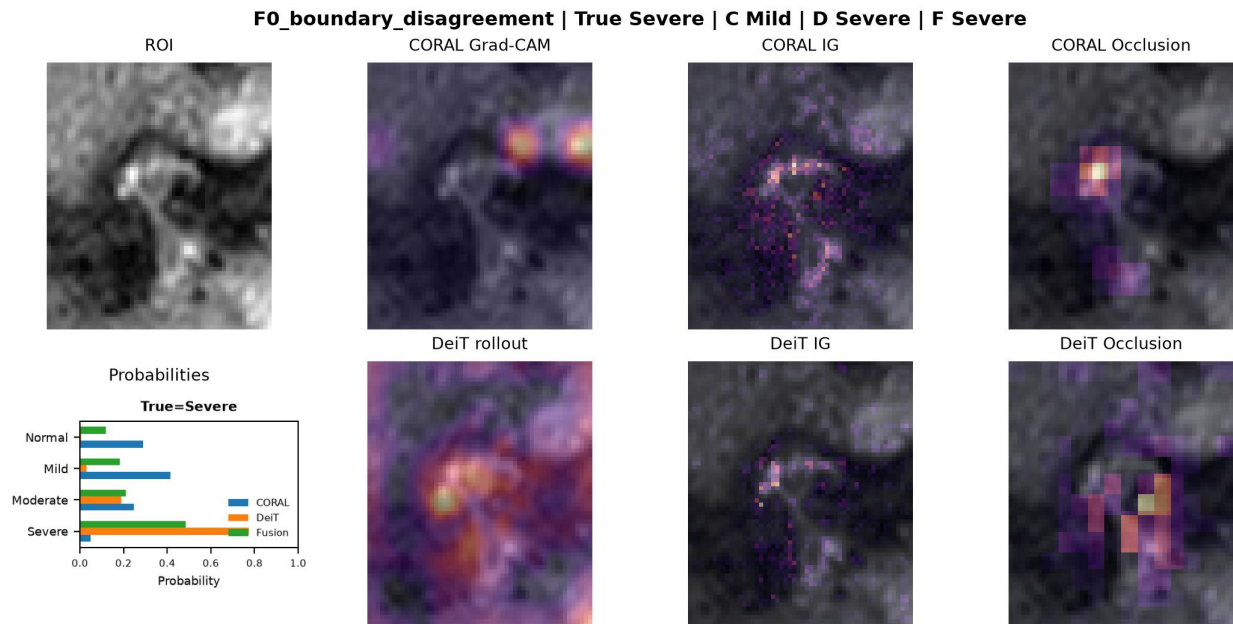
outer_fold	role	case	y_true	coral_pred	deit_pred	fusion_pred	paper_display_set	patient_id
0	benefit_correction	F0_benefit_correction	1	0	2	1	Main6	443
0	boundary_disagreement	F0_boundary_disagreement	3	1	3	3	Main6	75
0	challenge_error	F0_challenge_error	3	0	1	0	Supplement	77
0	correct_agreement	F0_correct_agreement	0	0	0	0	Supplement	4
1	benefit_correction	F1_benefit_correction	1	0	2	1	Supplement	38
1	boundary_disagreement	F1_boundary_disagreement	2	0	3	3	Supplement	111
1	challenge_error	F1_challenge_error	3	0	1	0	Main6	60
1	correct_agreement	F1_correct_agreement	0	0	0	0	Supplement	16
2	benefit_correction	F2_benefit_correction	2	1	3	2	Supplement	76
2	boundary_disagreement	F2_boundary_disagreement	3	0	2	2	Supplement	304
2	challenge_error	F2_challenge_error	3	0	1	1	Supplement	357
2	correct_agreement	F2_correct_agreement	0	0	0	0	Main6	6
3	benefit_correction	F3_benefit_correction	3	2	3	3	Supplement	209
3	boundary_disagreement	F3_boundary_disagreement	2	0	2	1	Main6	137
3	challenge_error	F3_challenge_error	2	0	0	0	Supplement	18
3	correct_agreement	F3_correct_agreement	0	0	0	0	Supplement	1
4	benefit_correction	F4_benefit_correction	1	0	2	1	Main6	483
4	boundary_disagreement	F4_boundary_disagreement	0	0	3	3	Supplement	241
4	challenge_error	F4_challenge_error	3	1	1	1	Supplement	337
4	correct_agreement	F4_correct_agreement	0	0	0	0	Supplement	8

Supplementary Figures S1-S20 use the frozen extended interpretability panels. Branch-specific displays are descriptive only; near-zero maps, where present, are retained as generated. No synthetic fused saliency or post hoc case replacement was used.

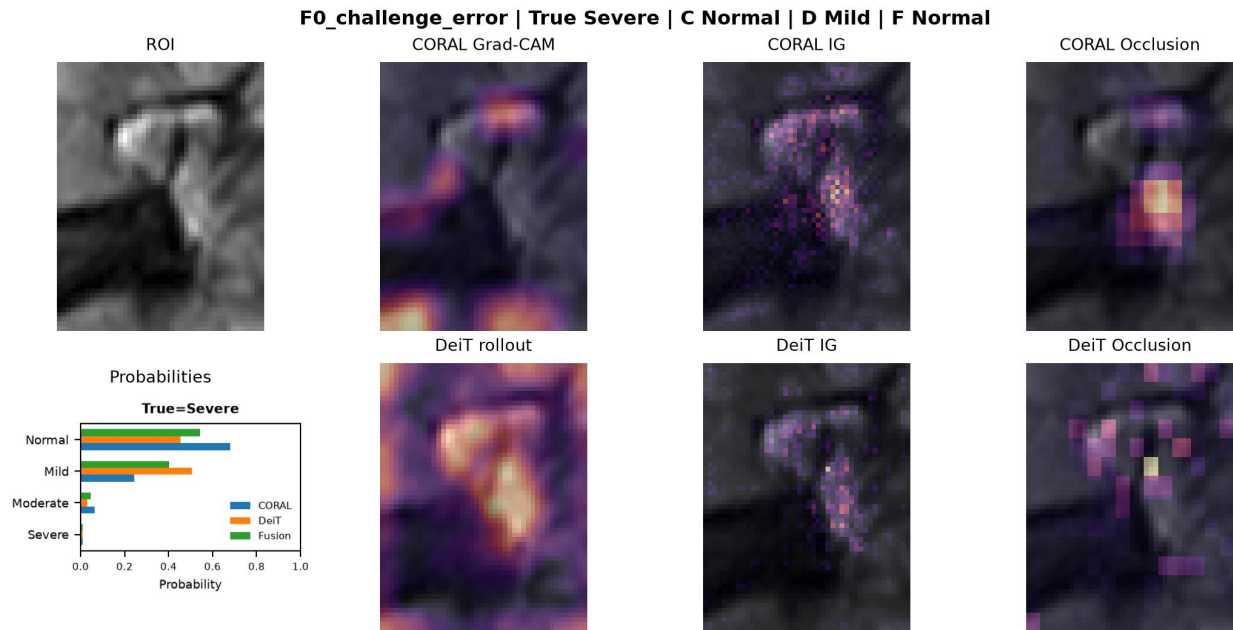
Supplementary Figure S1. Locked interpretability case - F0 benefit correction



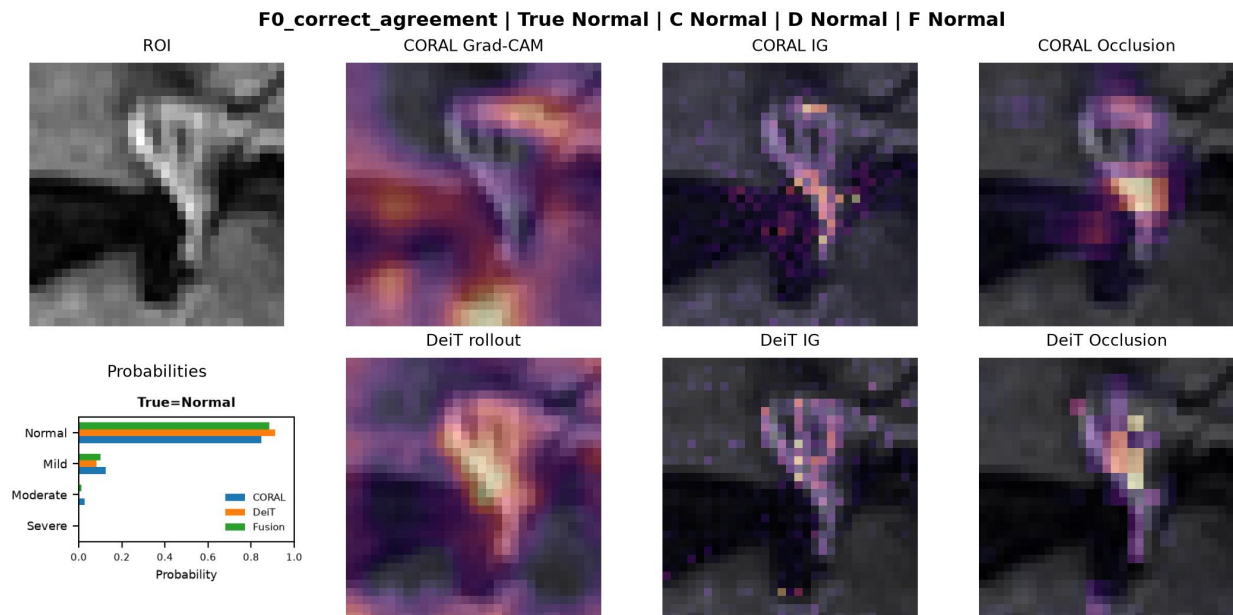
Supplementary Figure S2. Locked interpretability case - F0 boundary disagreement



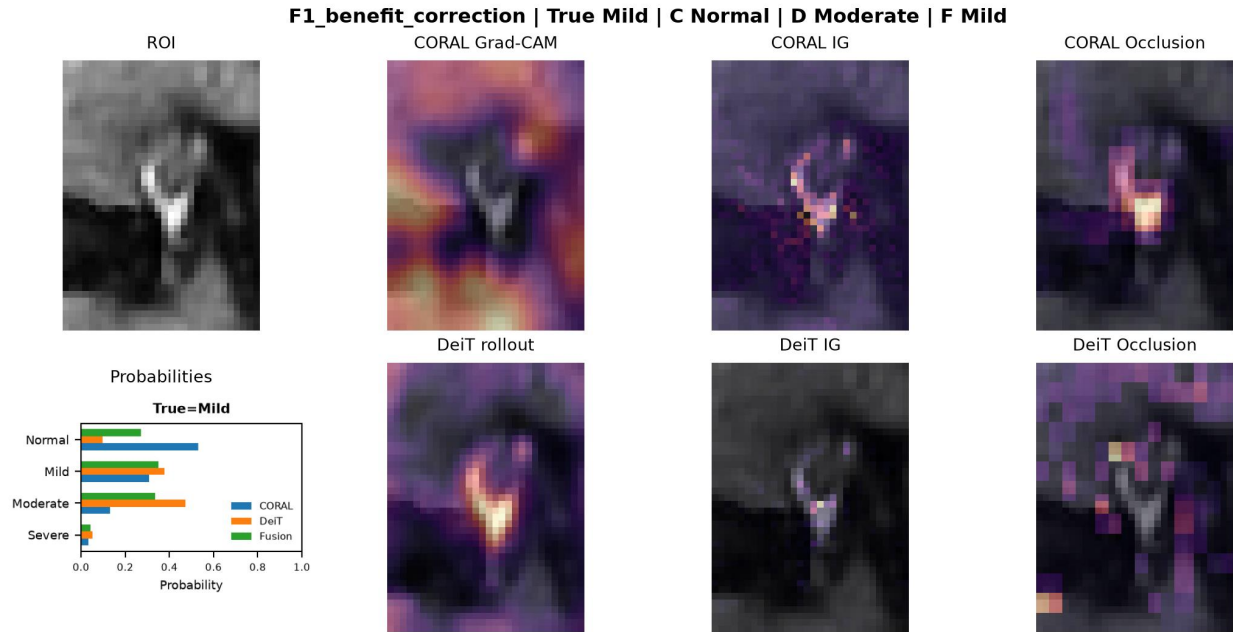
Supplementary Figure S3. Locked interpretability case - F0 challenge error



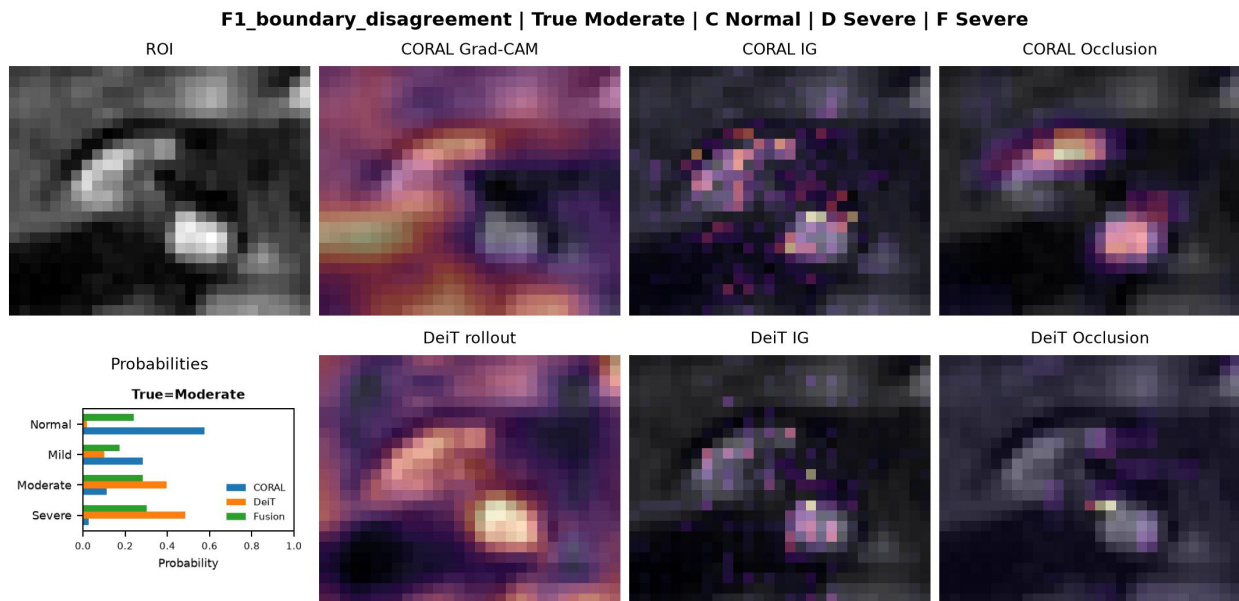
Supplementary Figure S4. Locked interpretability case - F0 correct agreement



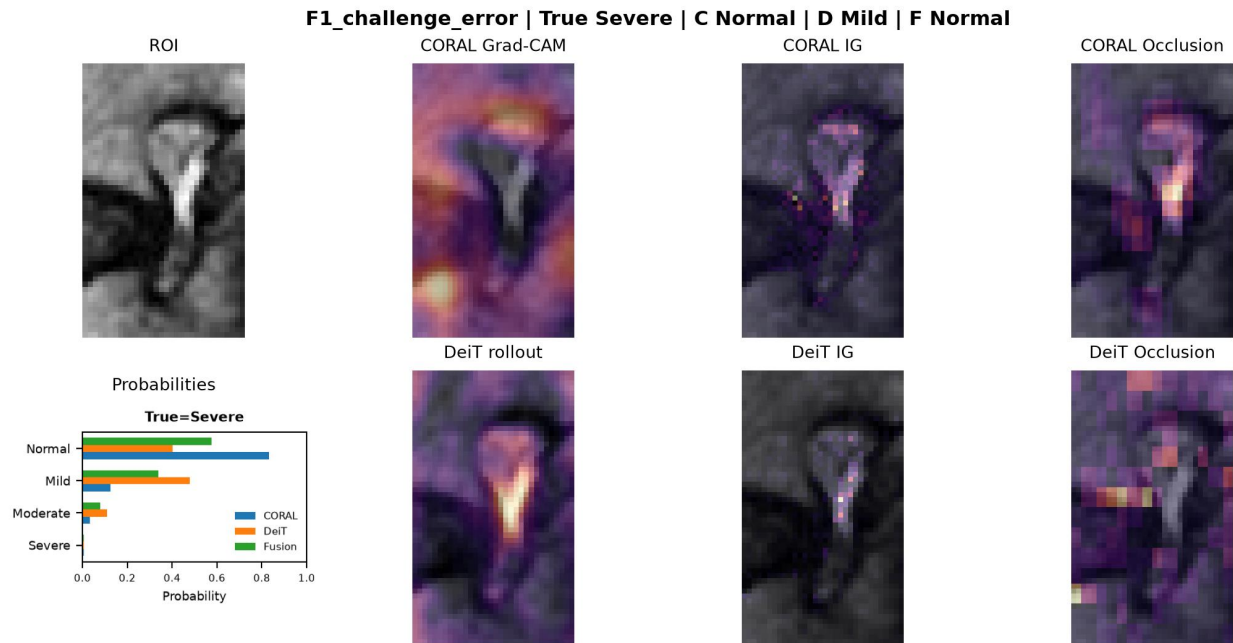
Supplementary Figure S5. Locked interpretability case - F1 benefit correction



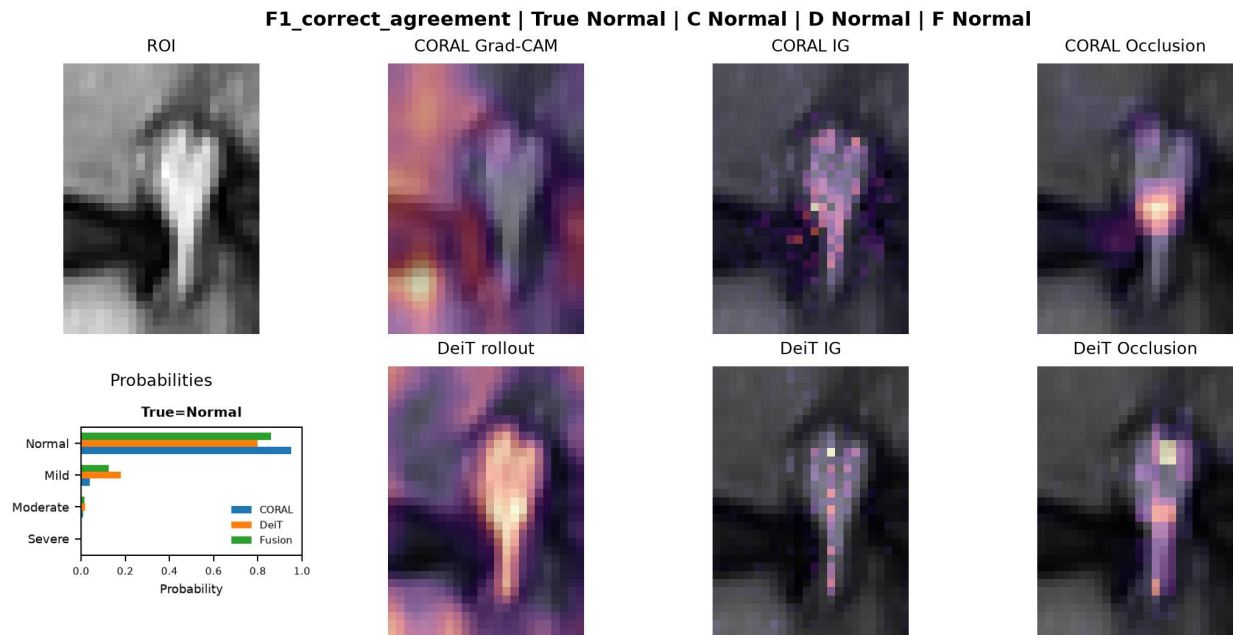
Supplementary Figure S6. Locked interpretability case - F1 boundary disagreement



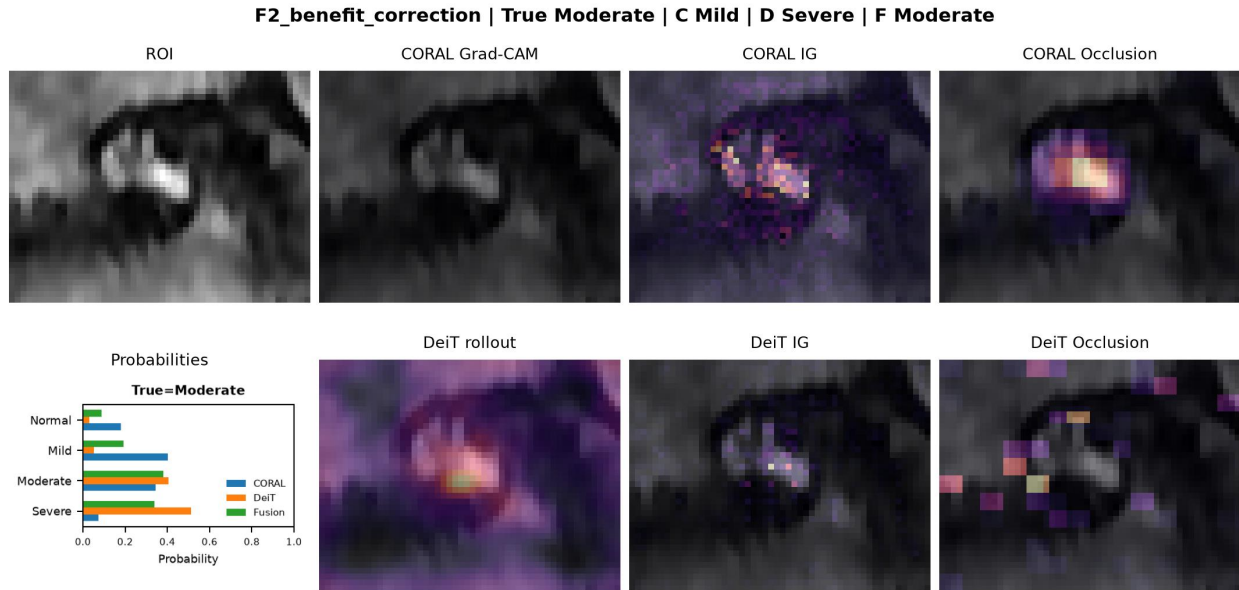
Supplementary Figure S7. Locked interpretability case - F1 challenge error



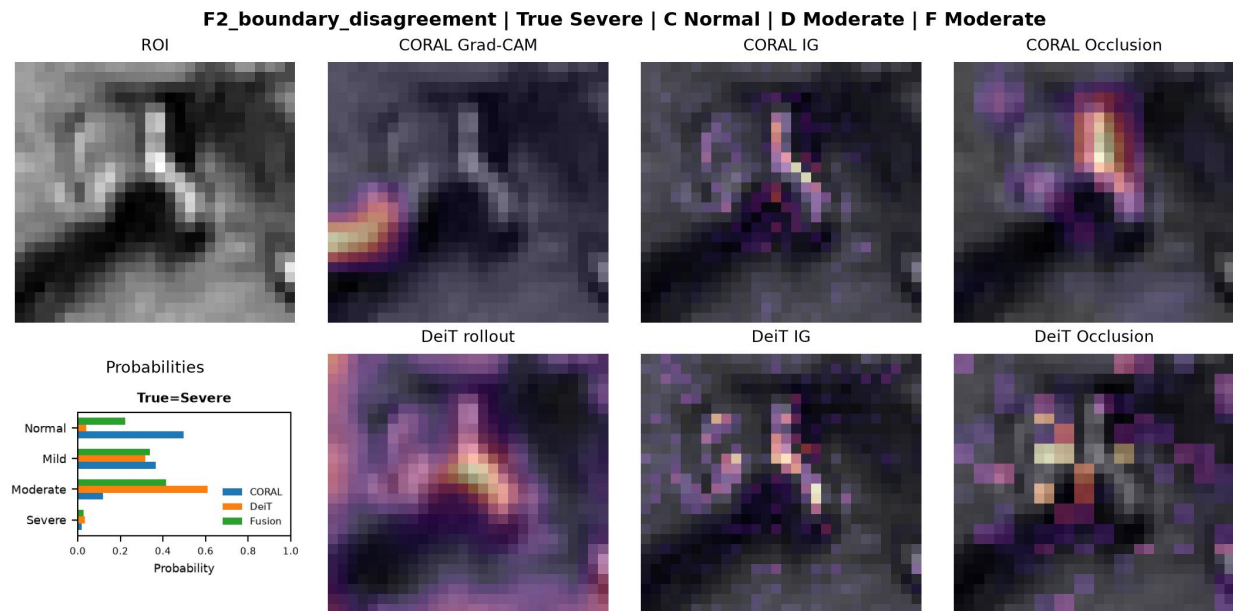
Supplementary Figure S8. Locked interpretability case - F1 correct agreement



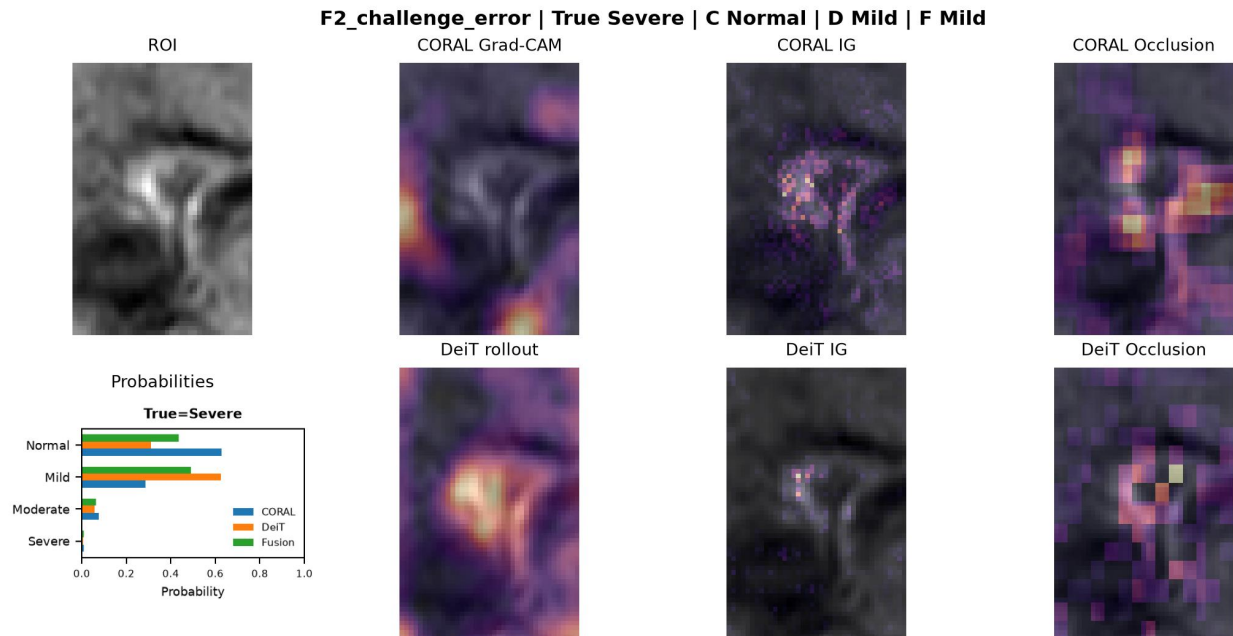
Supplementary Figure S9. Locked interpretability case - F2 benefit correction



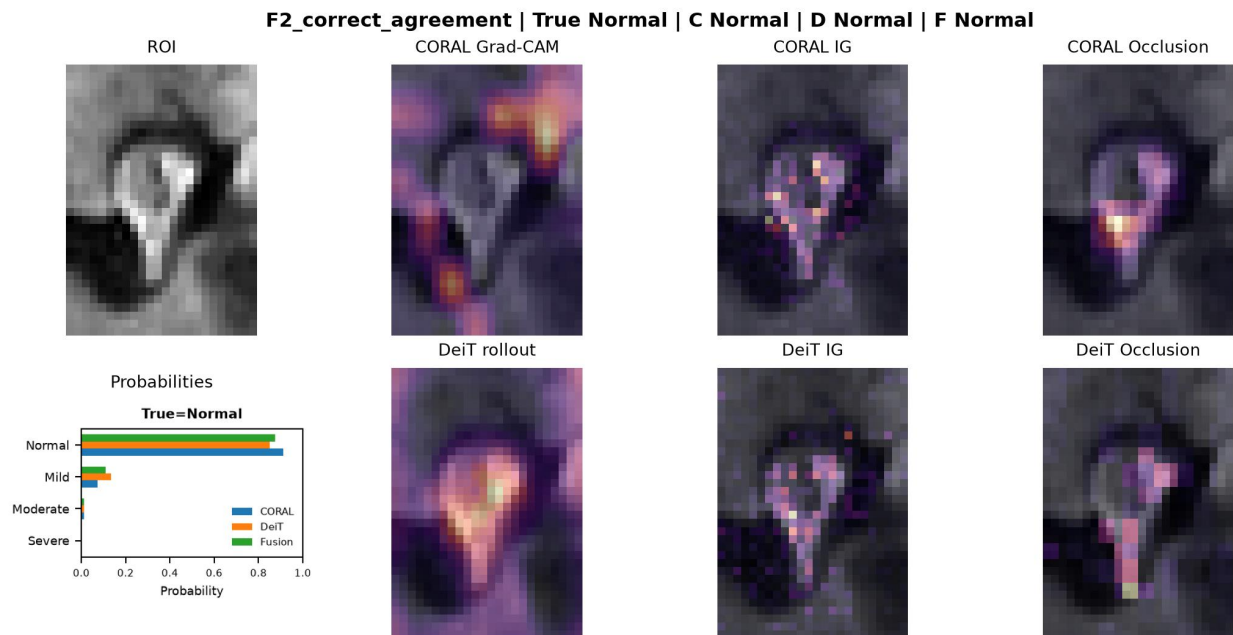
Supplementary Figure S10. Locked interpretability case - F2 boundary disagreement



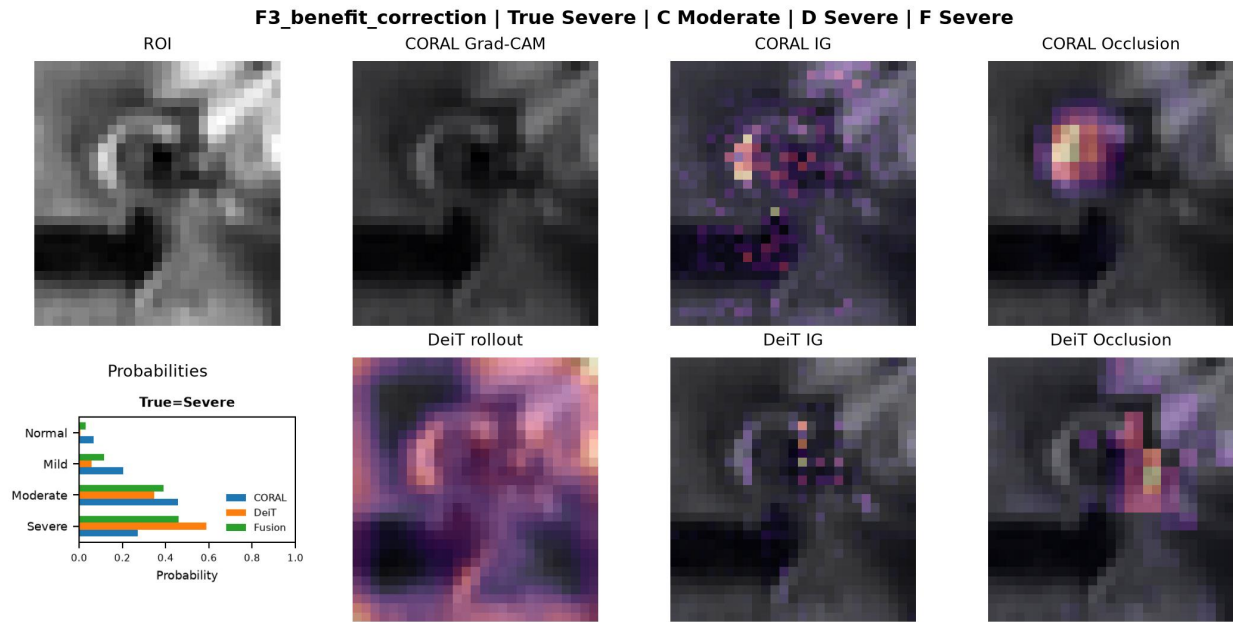
Supplementary Figure S11. Locked interpretability case - F2 challenge error



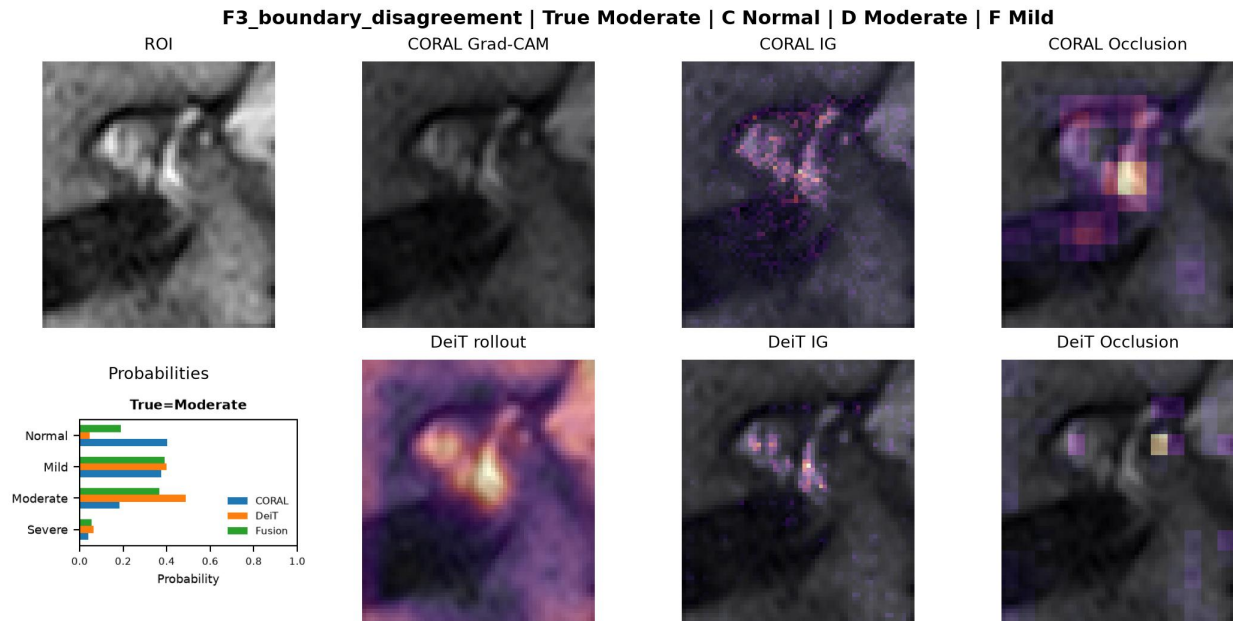
Supplementary Figure S12. Locked interpretability case - F2 correct agreement



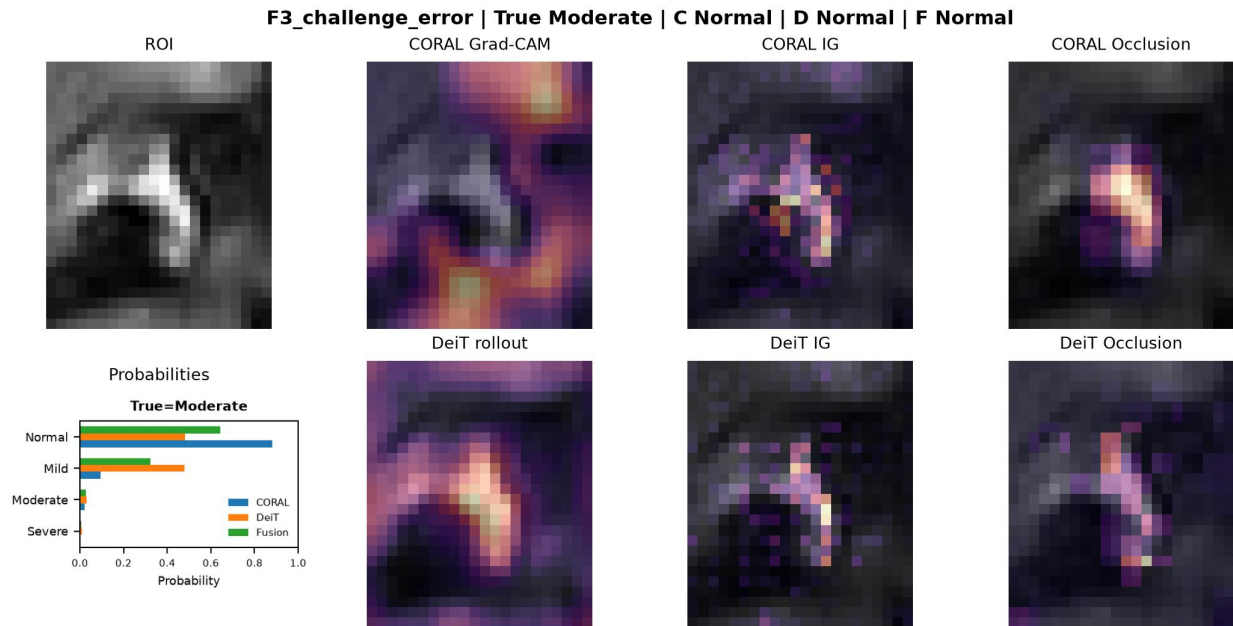
Supplementary Figure S13. Locked interpretability case - F3 benefit correction



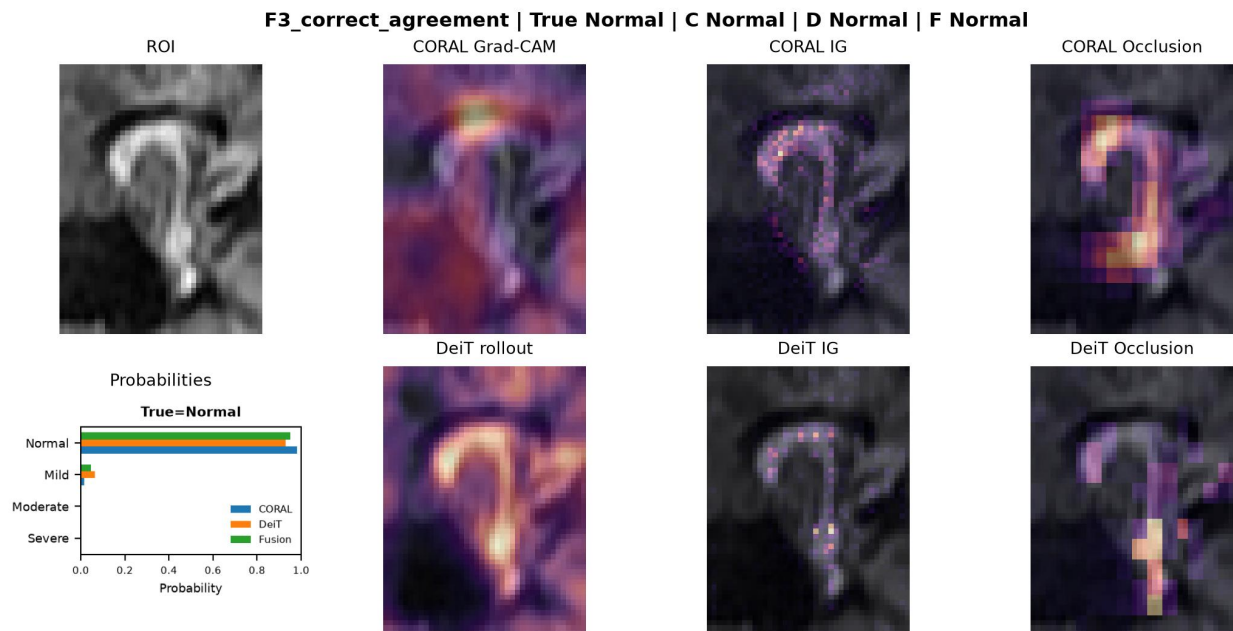
Supplementary Figure S14. Locked interpretability case - F3 boundary disagreement



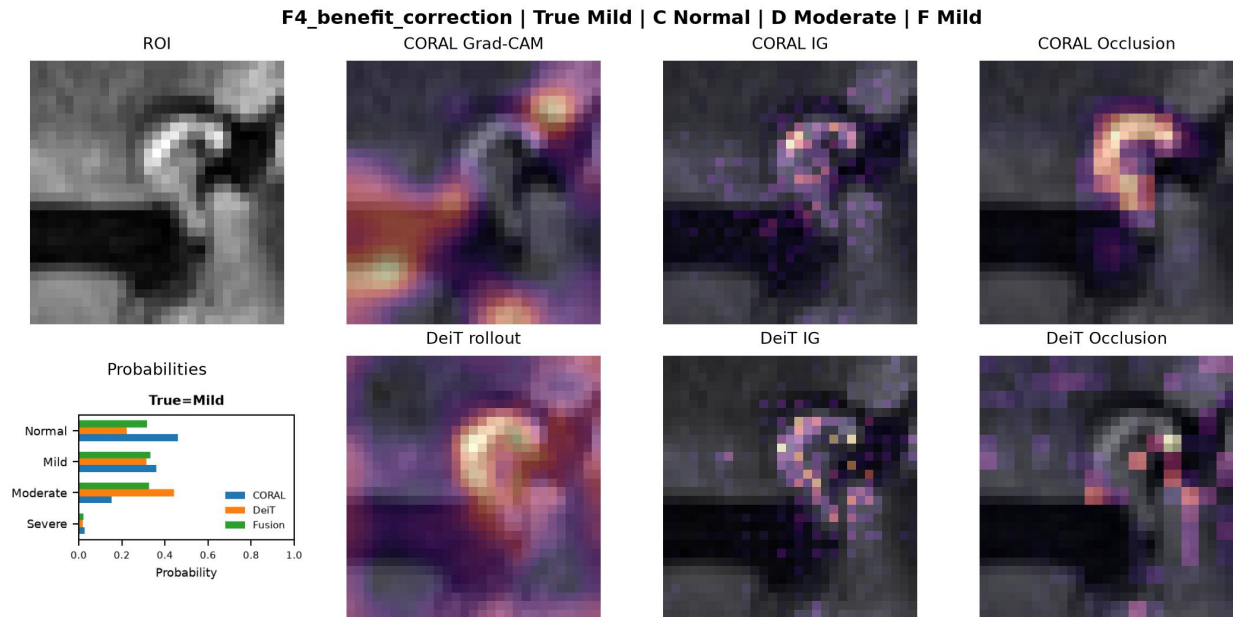
Supplementary Figure S15. Locked interpretability case - F3 challenge error



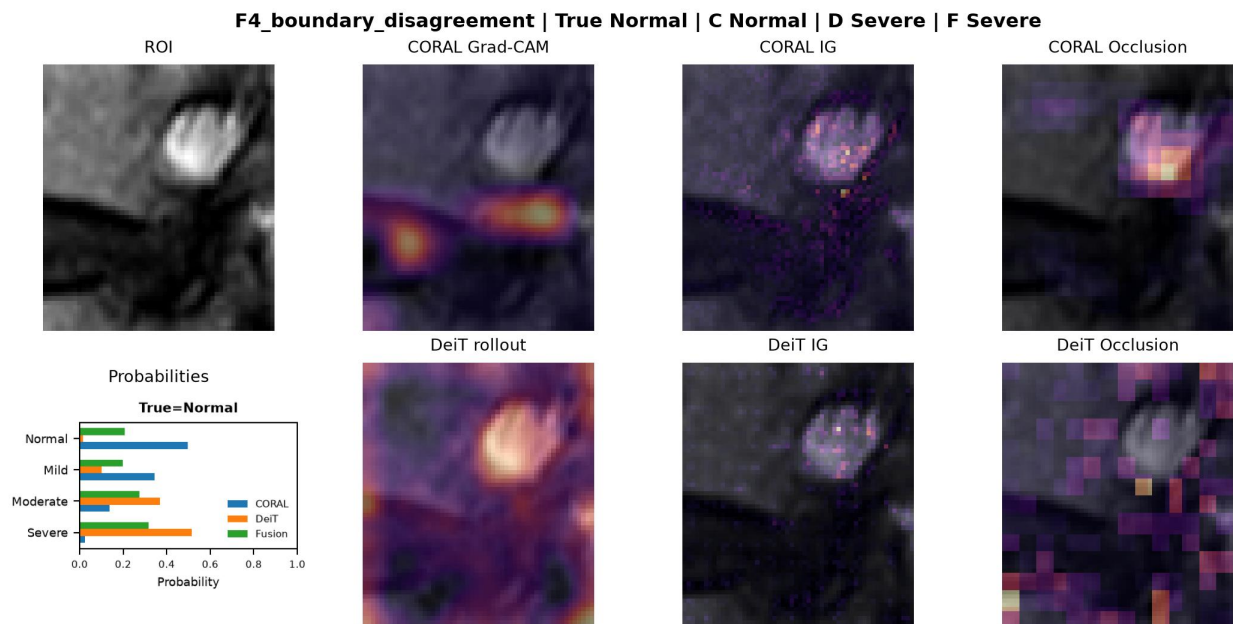
Supplementary Figure S16. Locked interpretability case - F3 correct agreement



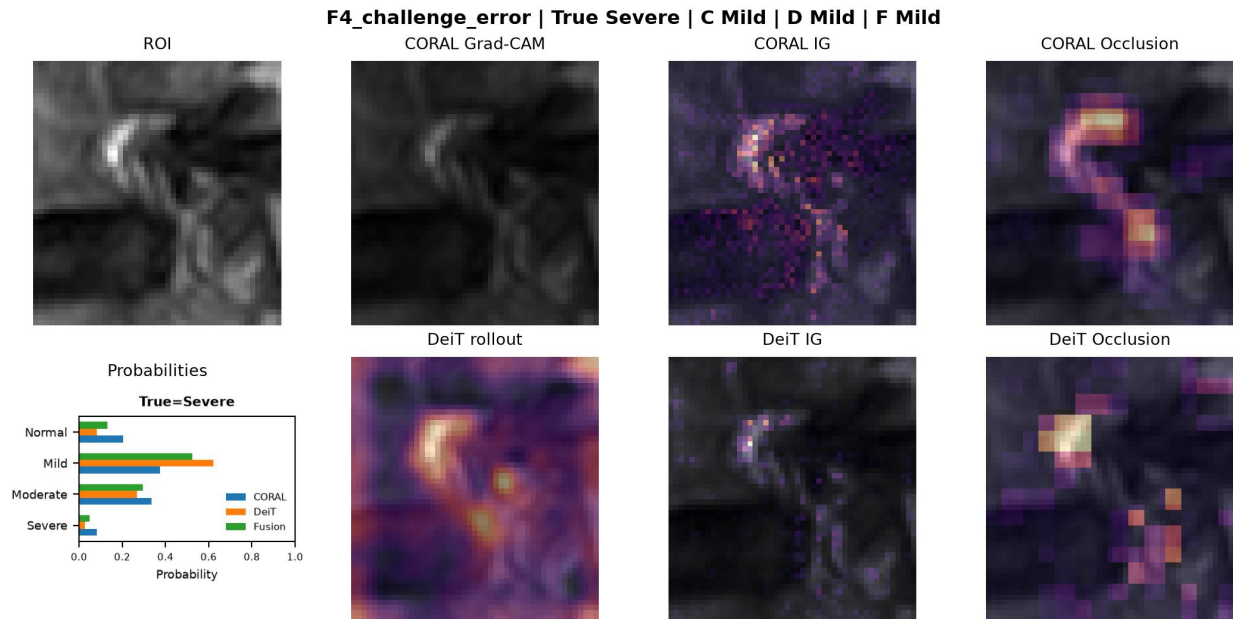
Supplementary Figure S17. Locked interpretability case - F4 benefit correction



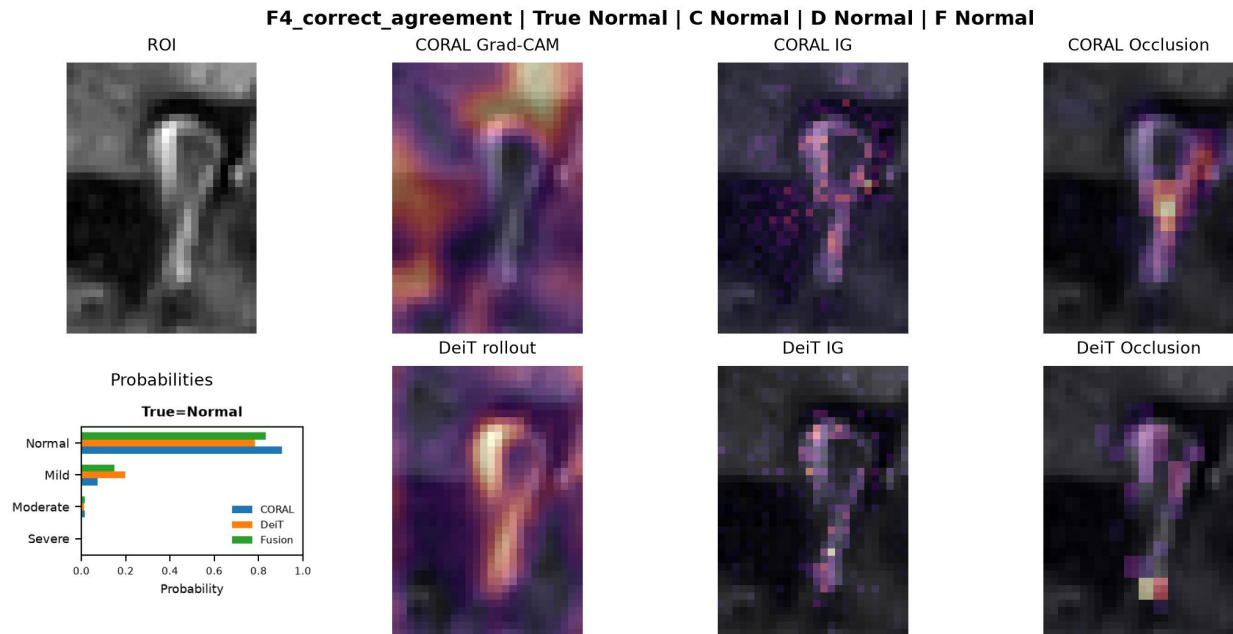
Supplementary Figure S18. Locked interpretability case - F4 boundary disagreement



Supplementary Figure S19. Locked interpretability case - F4 challenge error



Supplementary Figure S20. Locked interpretability case - F4 correct agreement



Supplementary Interpretability S2. Deletion faithfulness

Deletion faithfulness used 320 frozen case/branch/method/masking-fraction measurements as a descriptive sanity check.

Positive random-minus-guided AUC means that saliency-guided masking reduced target-class probability more strongly than random masking. This analysis is not a population-level inferential endpoint and does not establish causal anatomy.

Supplementary Table S8. Frozen random-minus-guided deletion AUC summary.

Branch	Method	Random-minus-guided AUC
CORAL-MSaux	Integrated gradients	0.1186
CORAL-MSaux	Occlusion	0.1198
Original DeiT	Integrated gradients	0.0300
Original DeiT	Occlusion	0.0666

Supplementary Reproducibility S1. Frozen-source provenance and exact checkpoint identities

Supplementary Table S9. Exact checkpoint SHA reverification used for the locked interpretability audit.

Model	Fold	Checkpoint file	SHA256	Status
CORAL-MSaux	0	best_val_ordinal_loss.pth	83de5d38bd2228c45def632c3321381e 2c4fd747e6411073360a9090fae28e22	PASS
Original DeiT	0	best_val_loss.pth	ac17fc555a07b6d0c00b2b56cf384b3c 8d8cbf11bb452654b52d98df13e67956	PASS
CORAL-MSaux	1	best_val_ordinal_loss.pth	9ea808b40ba4898d85d465752de26f1b 916e549f2fa91b0101b54997d61fa69c	PASS
Original DeiT	1	best_val_loss.pth	5a1e0e6be8b86980927afd015dbfa087 e28b7726882e5802a9a9da7e490066ac	PASS
CORAL-MSaux	2	best_val_ordinal_loss.pth	67ae1e0d8e0a20de9bc95b660701d507 a5f19060920f9d9e579d677d5478c434	PASS
Original DeiT	2	best_val_loss.pth	c1a4f872e5b09304cf4679fe88449bb0 f2b3a3d993fad98d07e802c977613ea8	PASS
CORAL-MSaux	3	best_val_ordinal_loss.pth	cbeffb78aaf6f07ab6dba86abfc334c3f 0c653ea63db16aae61e83d85aaa67a7a	PASS
Original DeiT	3	best_val_loss.pth	23e734e60579870fcce261c6e901d454 2f4a391266f61cf103d7863e3245b461	PASS
CORAL-MSaux	4	best_val_ordinal_loss.pth	d98a0e091dec42c9d3602e393aeb6388 1fc23133247e549bbf7633c8a5def102	PASS
Original DeiT	4	best_val_loss.pth	880a10df50ee64266714979b229b5bca fdb3dc76bd4e363aaf6f4a43145de05c	PASS

Supplementary Table S10. Key frozen source identifiers and SHA256 provenance.

Asset	Identifier	SHA256	Rows
Outer-fold reference map	Authoritative_Reference_OuterFold_Map_2978.csv	68262ebb6a8c06704209c565c3aad217 f17dc492ffc4fdd356e15ee463067569	2978
14-system reader-facing system registry	System_Registry.csv	12ceb8ebb5f022349aae6012577c1585e7138096a56d0be11bbf15550a77f862	14

Additional file 1 - Supplementary Material

Asset	Identifier	SHA256	Rows
Canonical cohort manifest	roi_splits.csv	2099a2a1b358ad3ee88d056064af289e adb59e038a476c7294f4be32e3b72b97	2978
20k paired-bootstrap reader-facing summary	Figure5_20kBootstrap.csv	aaedc239b3ee6b73a4c7bfc2603f3bd258b5cc488df511bf59f72ff86acec3cd	30
Modern-arm paper handoff	ModernArm_Final_Paper_Handoff.md	eae9d3a9d215a3d21c6377684168614d f108c83e90c0f000ce058460daf52ccd	
20-case interpretability metadata	Step06A9B_ExtendedInterpretability_Metadata_20Cases.csv	477ca156dddb7c2a400a66e4d01f89e5 ee60eaac013422f25193df9d4131d4c6	20
Deletion-faithfulness source	Step06A9B_FaithfulnessDeletion_Long.csv	1b9d0e22ffc8d1e7f1375e074f3e2622 4ae1e1853e2167d1767242702b2fb9c6	320
Checkpoint reverification table	Step06A9B_ExactCheckpointReverification.csv	47e7a5fa69571bc1405665e0c9aba2a1 d797fc90d9768af8a59e826a92b0cf07	10
Extended interpretability set	per_case_extended	set of 20 frozen panels	20 panels
Raw interpretability maps	raw_maps	set of 20 frozen NPZ archives	20 archives

Absolute workstation/server paths are omitted from reader-facing material. Full path-level provenance is retained in the internal Project Control registry and reproducibility package.

Scope boundary

This Additional file 1 contains only frozen evidence supporting the final 14-system paper. No scientific result in the frozen main manuscript was changed.