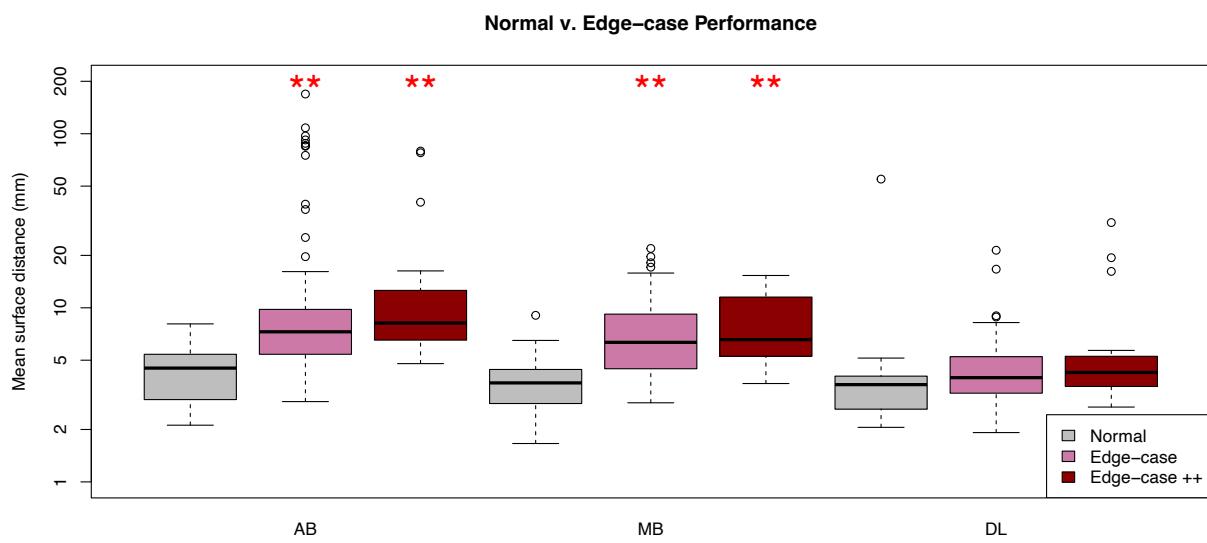
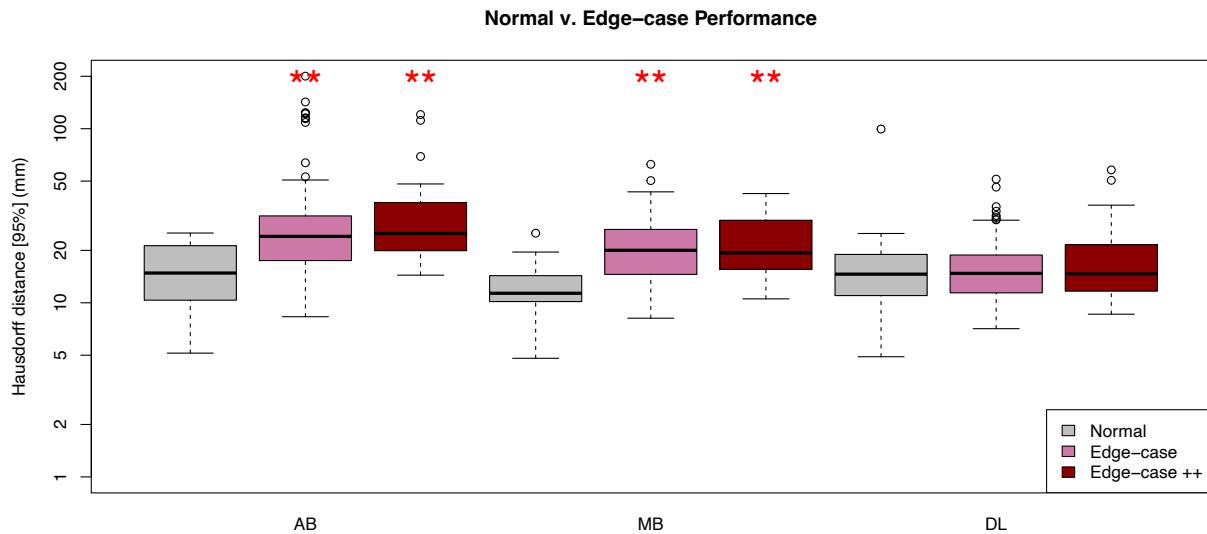


1 **Supplementary Figure 1:** Overall performance of autosegmentation tools on normal and edge
2 case cohorts. In the upper and lower panels, the distribution of 95% Hausdorff distances and
3 mean surface distances, respectively, in millimeters (y-axis) is shown as box plots for each of
4 three cohorts of individuals (“Normal” shown in gray, “Edge-case” with a single anatomic
5 variant shown in pink, and “Edge-case ++” with two or more simultaneous anatomic variants
6 shown in dark red), where each datapoint is an average across all structures for that individual.
7 Performance is reported for each of three autosegmentation tools: atlas-based
8 autosegmentation [AB], model-based segmentation [MB], and deep-learning based
9 segmentation [DL]). Statistically significant differences between normal and edge case
10 performance are denoted by asterisks, where (*) and (**) represent $p < 0.05$ and $p < 0.001$,
11 respectively (Wilcoxon Rank-Sum test).

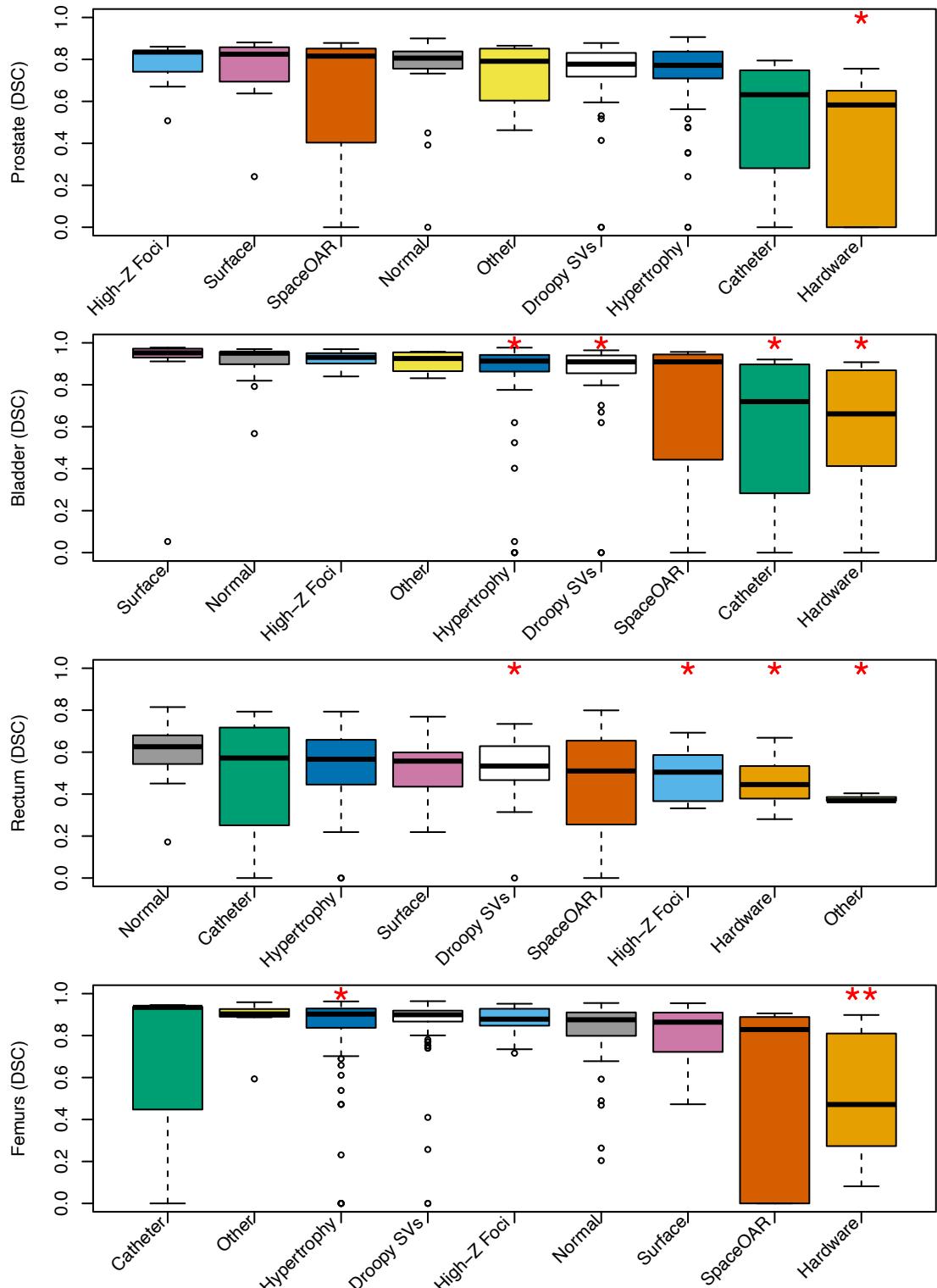
12

13



15 **Supplementary Figure 2:** Structure-specific performance of deep learning autosegmentation
16 across normal and discrete edge case cohorts. The distribution of Dice similarity coefficients
17 (DSC, y-axis) is shown here as box plots for each of nine cohorts of individuals corresponding to
18 “Normal” (gray), and eight distinct classes of anatomical variation defined and numbered in
19 Methods: 1) “Hypertrophy” (blue), 2) “Droopy SVs” (white), 3) “Hardware” (light orange), 4)
20 “Surface” (pink), 5) “High-Z Foci” (light blue), 6) “Catheter” (green), 7) “SpaceOAR” (orange),
21 and 8) “Other” (yellow), sorted according to median DSC. Each pane corresponds to a specific
22 structure, as labeled along y-axis, with “Femurs” representing the average result from bilateral
23 structures. Statistically significant differences between normal and edge case performance are
24 denoted by asterisks, where (*) and (**) represent $p < 0.05$ and $p < 0.001$, respectively (Wilcoxon
25 Rank-Sum test).

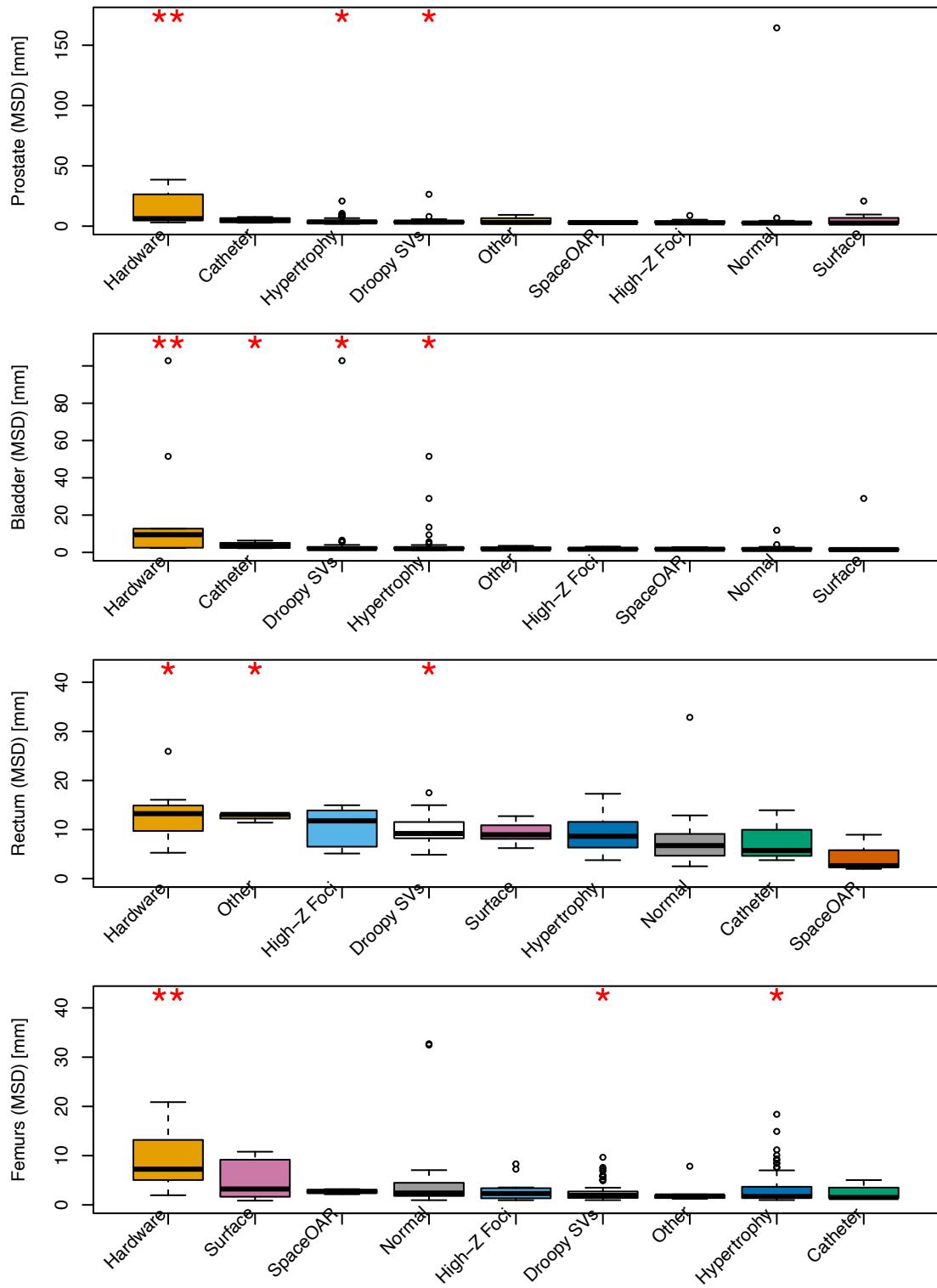
26



28 **Supplementary Figure 3:** Structure-specific performance of deep learning autosegmentation
29 across normal and discrete edge case cohorts. The distribution of mean surface distances in
30 millimeters (MSD, y-axis) is shown here as box plots for each of nine cohorts of individuals
31 corresponding to “Normal” (gray), and eight distinct classes of anatomical variation defined and
32 numbered in Methods: 1) “Hypertrophy” (blue), 2) “Droopy SVs” (white), 3) “Hardware” (light
33 orange), 4) “Surface” (pink), 5) “High-Z Foci” (light blue), 6) “Catheter” (green), 7) “SpaceOAR”
34 (orange), and 8) “Other” (yellow), sorted according to median MSD. Each pane corresponds to a
35 specific structure, as labeled along y-axis, with “Femurs” representing the average result from
36 bilateral structures. Statistically significant differences between normal and edge case
37 performance are denoted by asterisks, where (*) and (**) represent $p < 0.05$ and $p < 0.001$,
38 respectively (Wilcoxon Rank-Sum test).

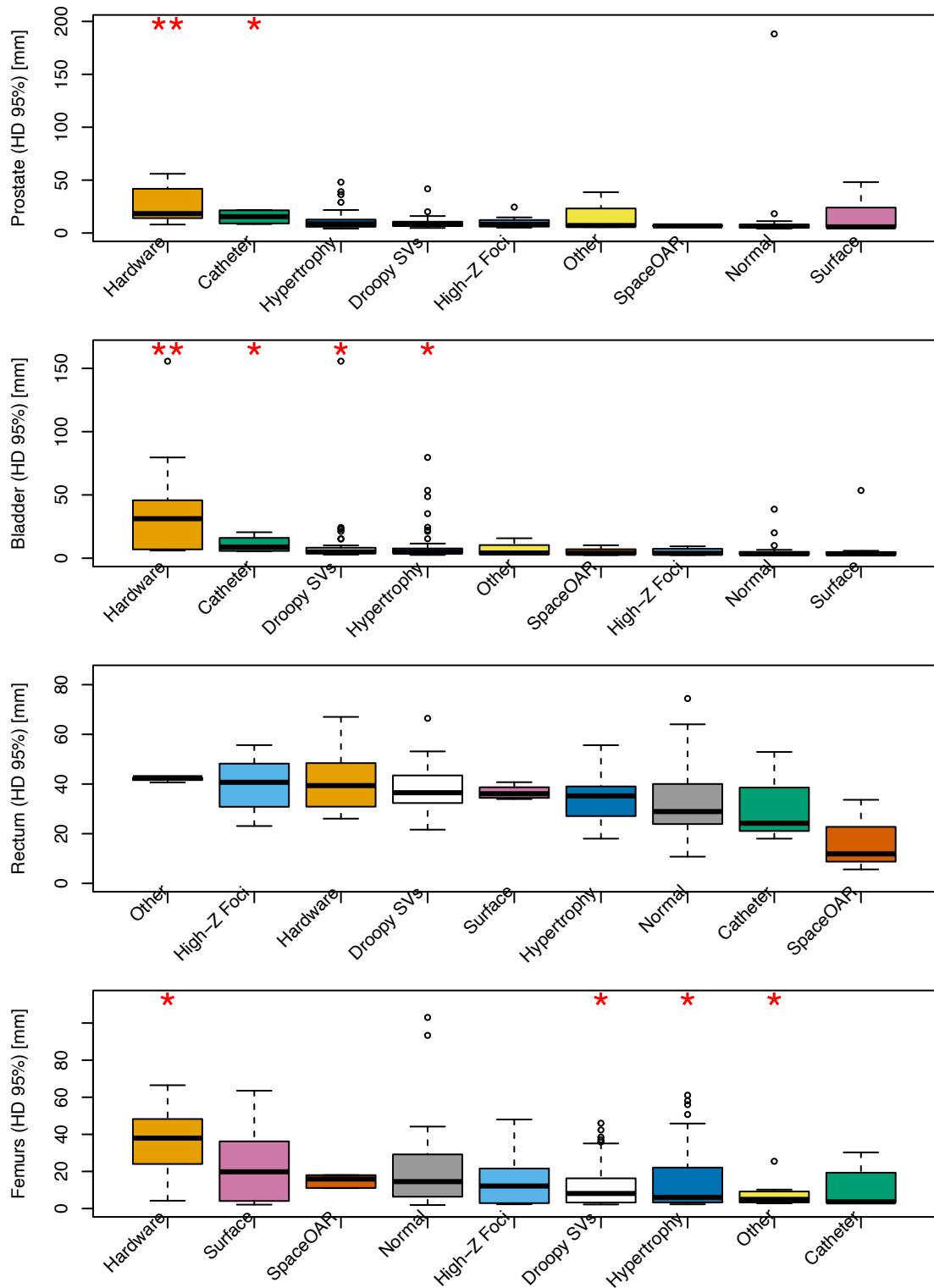
39

40



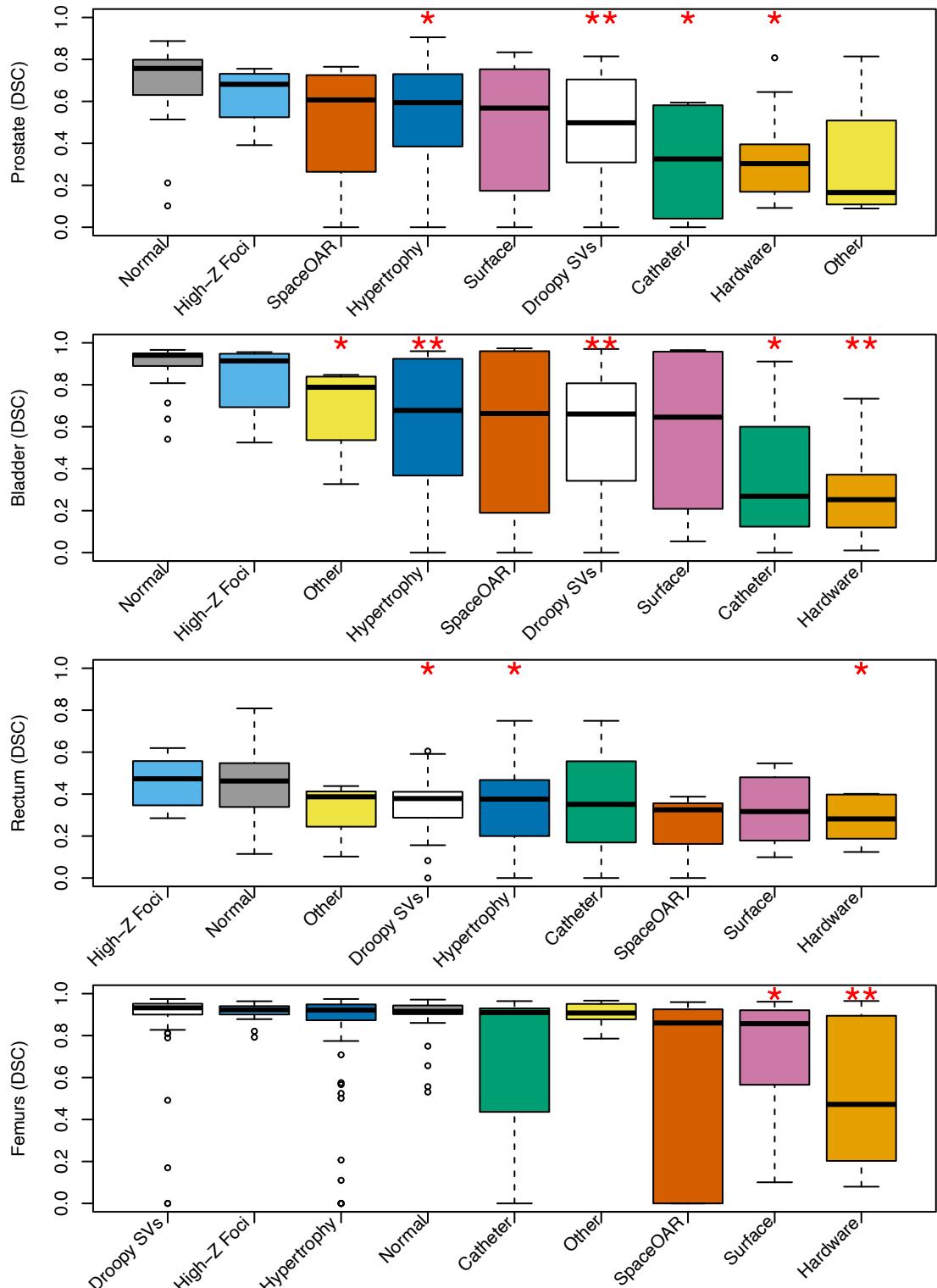
43 **Supplementary Figure 4:** Structure-specific performance of deep learning autosegmentation
44 across normal and discrete edge case cohorts. The distribution of 95% Hausdorff distances in
45 millimeters (HD 95%, y-axis) is shown here as box plots for each of nine cohorts of individuals
46 corresponding to “Normal” (gray), and eight distinct classes of anatomical variation defined and
47 numbered in Methods: 1) “Hypertrophy” (blue), 2) “Droopy SVs” (white), 3) “Hardware” (light
48 orange), 4) “Surface” (pink), 5) “High-Z Foci” (light blue), 6) “Catheter” (green), 7) “SpaceOAR”
49 (orange), and 8) “Other” (yellow), sorted according to median HD 95%. Each pane corresponds
50 to a specific structure, as labeled along y-axis, with “Femurs” representing the average result
51 from bilateral structures. Statistically significant differences between normal and edge case
52 performance are denoted by asterisks, where (*) and (**) represent $p < 0.05$ and $p < 0.001$,
53 respectively (Wilcoxon Rank-Sum test).

54

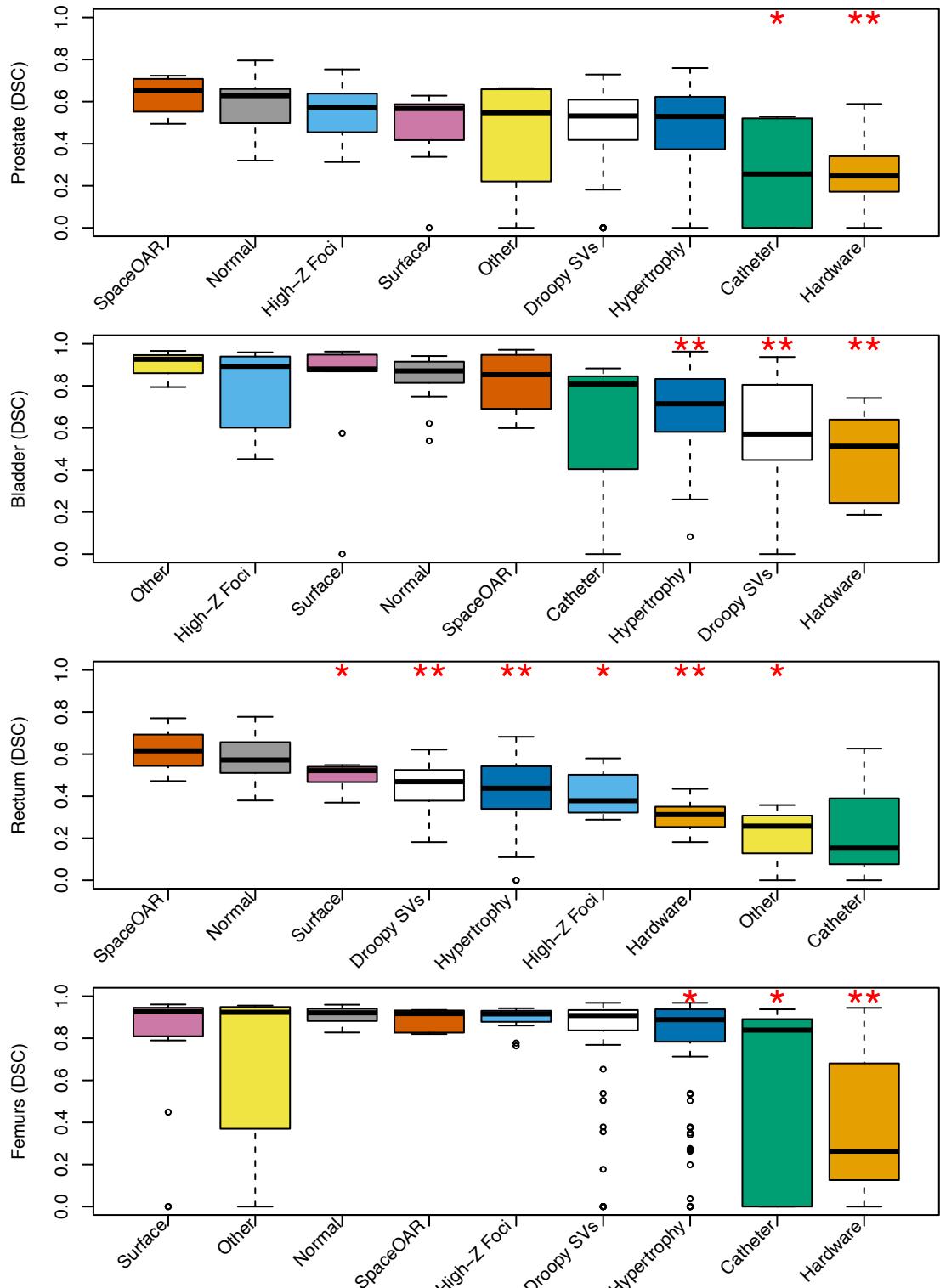


56 **Supplementary Figure 5:** Structure-specific performance of model-based autosegmentation
57 across normal and discrete edge case cohorts. The distribution of Dice similarity coefficients
58 (DSC, y-axis) is shown here as box plots for each of nine cohorts of individuals corresponding to
59 “Normal” (gray), and eight distinct classes of anatomical variation defined and numbered in
60 Methods: 1) “Hypertrophy” (blue), 2) “Droopy SVs” (white), 3) “Hardware” (light orange), 4)
61 “Surface” (pink), 5) “High-Z Foci” (light blue), 6) “Catheter” (green), 7) “SpaceOAR” (orange),
62 and 8) “Other” (yellow), sorted according to median DSC. Each pane corresponds to a specific
63 structure, as labeled along y-axis, with “Femurs” representing the average result from bilateral
64 structures. Statistically significant differences between normal and edge case performance are
65 denoted by asterisks, where (*) and (**) represent $p < 0.05$ and $p < 0.001$, respectively (Wilcoxon
66 Rank-Sum test).

67

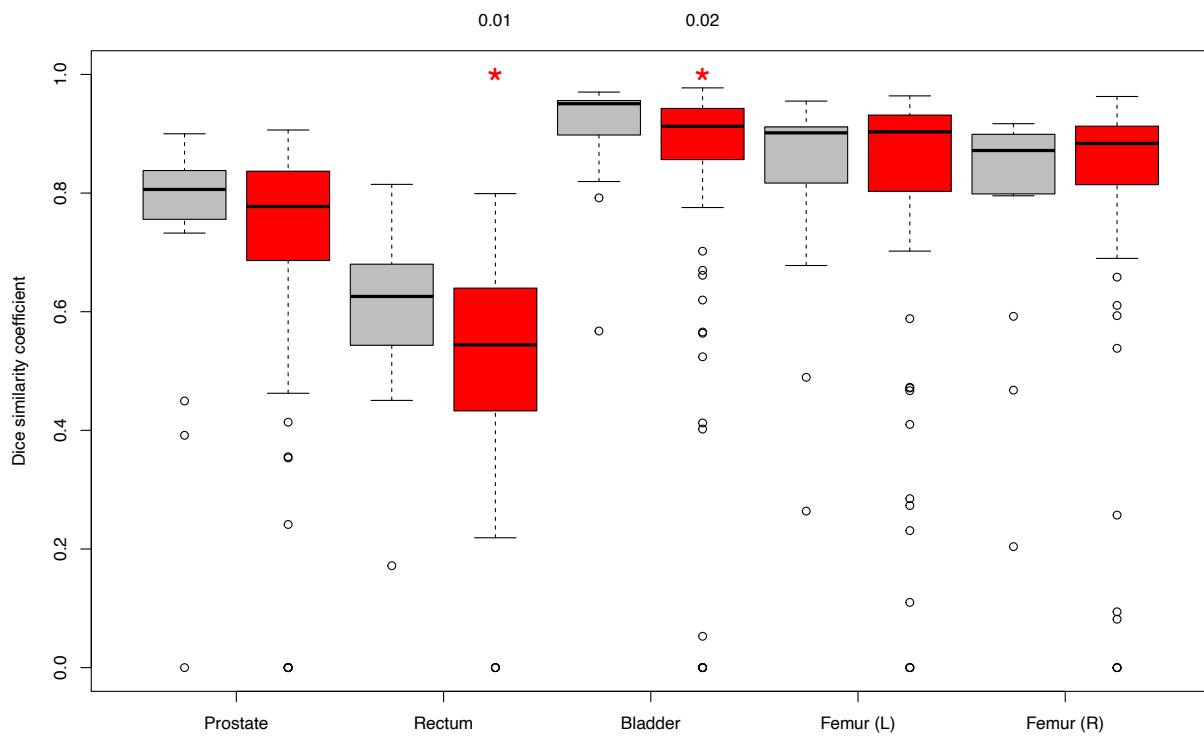


69 **Supplementary Figure 6:** Structure-specific performance of atlas-based autosegmentation
70 across normal and discrete edge case cohorts. The distribution of Dice similarity coefficients
71 (DSC, y-axis) is shown here as box plots for each of nine cohorts of individuals corresponding to
72 “Normal” (gray), and eight distinct classes of anatomical variation defined and numbered in
73 Methods: 1) “Hypertrophy” (blue), 2) “Droopy SVs” (white), 3) “Hardware” (light orange), 4)
74 “Surface” (pink), 5) “High-Z Foci” (light blue), 6) “Catheter” (green), 7) “SpaceOAR” (orange),
75 and 8) “Other” (yellow), sorted according to median DSC. Each pane corresponds to a specific
76 structure, as labeled along y-axis, with “Femurs” representing the average result from bilateral
77 structures. Statistically significant differences between normal and edge case performance are
78 denoted by asterisks, where (*) and (**) represent $p < 0.05$ and $p < 0.001$, respectively (Wilcoxon
79 Rank-Sum test).



82 **Supplementary Figure 7:** Structure-specific performance of deep learning autosegmentation
 83 across normal and edge case cohorts. The distribution of Dice similarity coefficients (DSC, y-
 84 axis) is shown here as box plots for each of five structures (“Prostate”, “Rectum”, “Bladder”,
 85 “Femur (L)”, and “Femur (R)”) corresponding to normal (gray) and edge cases (red). Statistically
 86 significant differences between normal and edge case performance are denoted by asterisks,
 87 where (*) represents $p < 0.05$ (Wilcoxon Rank-Sum test).

88



89

Supplementary Table 1. Characteristics of different classes of anatomical edge cases. P-values, calculated by Wilcoxon Rank-Sum test for each class of edge case v. normal, are shown where significant (<0.05). Volumes (vol.) in cubic centimeters (cm³) are median values per group.

Class	Description	Size (n)	Prostate vol. (cm ³)	pval (Prostate)	Bladder vol. (cm ³)	pval (Bladder)	Rectum vol. (cm ³)	pval (Rectum)
	Normal	19	50.91		177.3		45.31	
1	Prostate hypertrophy	52	80.46	<0.001	125.47		32.18	0.02
2	"Droopy" seminal vesicles	37	61.8	0.03	114.86	0.04	32.89	0.04
3	Hip arthroplasty	11	53.61		72.51	<0.01	32.37	0.02
4	Prostate surface irregularity	9	63.45		225.25		32.8	
5	Calcifications	8	61.17		113.91		24.84	0.01
6	Foley catheter	4	112.66	0.04	108.45		40.04	
7	SpaceOAR	4	56.72		173.63		62.51	
8	Narrow rectum	2	55.01		245.62		26.57	
8	In-field bowel	1	50.83		61.55		24.97	
8	Morbid obesity	1	88.43		180.12		51.52	
	TOTAL	131						

Supplementary Table 2. Cohort details (individual cases). Structure volumes listed in cubic centimeters (cc).

ID	Cohort	Cohort (2)	Prostate (cc)	Bladder (cc)	Rectum (cc)
1	Prostate surface irregularity		63.45	147.53	40.01
2	"Droopy" seminal vesicles		47.60	157.43	19.05
3	"Droopy" seminal vesicles		46.98	83.62	27.26
4	Prostate hypertrophy		137.13	282.27	53.48
5	Prostate hypertrophy		72.62	119.53	72.56
6	Seeds, calcifications		83.96	158.02	16.85
7	Prostate hypertrophy		206.97	178.54	32.48
8	Normal		128.95	127.63	42.04
9	Prostate surface irregularity		50.29	360.88	15.77
10	Hip arthroplasty		38.88	87.34	33.06
11	Seeds, calcifications		69.45	104.80	55.81
12	Hip arthroplasty		57.83	72.51	44.09
13	Prostate hypertrophy		65.92	92.93	25.85
14	Prostate surface irregularity	Prostate hypertrophy	194.03	257.90	32.80
15	Prostate hypertrophy		151.35	251.53	57.33
16	Prostate hypertrophy	"Droopy" seminal vesicles	56.67	127.16	36.72
17	"Droopy" seminal vesicles		61.80	115.00	21.41
18	"Droopy" seminal vesicles		56.99	152.49	67.82
19	Prostate hypertrophy		89.11	188.66	32.29
20	Prostate surface irregularity		67.89	165.18	32.05
21	Prostate hypertrophy		84.27	110.80	18.70
22	Foley catheter		79.69	79.79	52.84
23	Prostate hypertrophy		136.43	98.19	44.30
24	In-field bowel		50.83	61.55	24.97
25	"Droopy" seminal vesicles		105.35	117.51	25.03
26	"Droopy" seminal vesicles	Hip arthroplasty	37.30	84.86	33.53
27	"Droopy" seminal vesicles		46.82	62.41	22.00
28	Hip arthroplasty		85.05	145.89	27.59
29	Foley catheter	Prostate hypertrophy	182.71	137.11	65.07
30	Seeds, calcifications		83.76	140.00	26.80
31	"Droopy" seminal vesicles		118.70	158.30	28.79
32	Prostate hypertrophy		79.70	236.66	18.52
33	Prostate hypertrophy		43.90	175.42	19.53
34	Prostate hypertrophy		166.90	114.33	17.79
35	Hip arthroplasty		54.18	63.12	27.03
36	Prostate hypertrophy		197.61	49.63	30.54
37	Prostate hypertrophy		75.19	332.99	29.63
38	Prostate hypertrophy	"Droopy" seminal vesicles	50.53	204.27	60.23
39	Seeds, calcifications		39.44	70.50	26.16
40	"Droopy" seminal vesicles		85.56	60.61	17.36
41	Foley catheter		43.81	73.99	27.25
42	Seeds, calcifications		51.46	61.56	23.52
43	"Droopy" seminal vesicles		48.03	66.74	16.66
44	Narrow rectum		63.61	254.20	32.47
45	Prostate hypertrophy	"Droopy" seminal vesicles	80.98	176.19	26.01
46	Prostate hypertrophy		94.34	122.04	17.10
47	"Droopy" seminal vesicles		65.38	242.32	50.79
48	"Droopy" seminal vesicles		85.76	295.42	24.34
49	Hip arthroplasty		36.97	103.59	23.62
50	"Droopy" seminal vesicles		95.36	114.86	43.57
51	Prostate surface irregularity	Prostate hypertrophy	151.46	507.35	28.31
52	Prostate hypertrophy		79.80	483.27	43.63
53	Seeds, calcifications		34.73	122.26	18.42
54	"Droopy" seminal vesicles		40.56	52.72	40.11
55	Prostate surface irregularity		75.66	348.29	90.39
56	Prostate hypertrophy		106.95	234.02	32.96
57	Prostate hypertrophy		47.87	90.25	49.44
58	Prostate surface irregularity		63.31	115.29	27.73
59	"Droopy" seminal vesicles		45.17	222.83	23.23
60	Prostate hypertrophy		96.76	162.70	18.09
61	Prostate hypertrophy		72.00	133.58	27.03
62	Prostate hypertrophy		53.57	129.84	18.80
63	"Droopy" seminal vesicles		96.04	71.08	56.81
64	Normal		57.90	102.95	23.29
65	Prostate hypertrophy		111.19	433.62	28.72
66	Prostate hypertrophy		53.43	54.13	29.63
67	Prostate hypertrophy	"Droopy" seminal vesicles	80.77	101.61	36.33
68	Prostate hypertrophy	"Droopy" seminal vesicles	58.58	190.80	37.47
69	Prostate hypertrophy		92.16	174.26	39.31
70	"Droopy" seminal vesicles	Hip arthroplasty	47.16	61.65	36.02
71	Hip arthroplasty		90.71	44.90	22.70
72	"Droopy" seminal vesicles		52.89	78.49	44.91
73	Normal		119.76	91.47	60.34
74	Prostate hypertrophy	Hip arthroplasty	53.61	71.37	30.28
75	"Droopy" seminal vesicles		57.35	234.30	59.89
76	Prostate hypertrophy		61.98	117.05	20.50
77	Prostate hypertrophy	"Droopy" seminal vesicles	55.16	100.78	57.48
78	Prostate surface irregularity		37.60	113.00	344.90
79	Prostate hypertrophy	"Droopy" seminal vesicles	86.07	78.74	42.40
80	Prostate hypertrophy	"Droopy" seminal vesicles	46.06	68.74	23.37
81	"Droopy" seminal vesicles		66.34	78.58	27.51
82	Narrow rectum		46.42	237.03	20.67
83	"Droopy" seminal vesicles		75.60	251.43	32.94
84	Seeds, calcifications		61.97	105.69	29.04
85	Prostate hypertrophy		56.40	86.47	30.02
86	Prostate surface irregularity		49.62	225.22	48.15
87	"Droopy" seminal vesicles		56.68	156.66	21.77
88	Prostate hypertrophy		120.75	69.87	40.62
89	"Droopy" seminal vesicles		80.14	160.11	25.38
90	Prostate hypertrophy		80.14	74.78	34.53
91	Normal		35.49	177.41	28.45
92	"Droopy" seminal vesicles		79.89	83.01	32.89
93	Prostate hypertrophy		115.34	47.50	63.81
94	"Droopy" seminal vesicles		125.60	85.65	30.66
95	"Droopy" seminal vesicles	Hip arthroplasty	99.86	205.13	32.37
96	"Droopy" seminal vesicles		55.14	51.55	28.99
97	Seeds, calcifications		60.37	122.13	22.93
98	Prostate hypertrophy		125.48	192.96	32.06
99	Prostate hypertrophy		63.87	123.79	23.54
100	Prostate hypertrophy		42.35	86.08	23.18
101	"Droopy" seminal vesicles		63.97	249.39	283.74
102	Normal		54.55	44.09	28.60
103	Hip arthroplasty		43.66	53.83	37.27
104	Prostate hypertrophy	"Droopy" seminal vesicles	73.62	54.44	44.95
105	Prostate hypertrophy		51.00	74.35	27.98
106	Prostate hypertrophy		42.55	56.43	48.48
107	Prostate hypertrophy		65.90	171.13	23.05
108	Prostate hypertrophy		81.56	93.68	20.76
109	Normal		31.92	130.16	42.33
110	Normal		62.22	246.52	86.16
111	Normal		66.95	133.91	87.26
112	Normal		64.31	99.51	22.67
113	Normal		50.20	258.12	41.15
114	Normal		61.68	272.42	54.52
115	SpaceOAR		60.85	689.82	67.81
116	Normal		66.84	209.13	53.23
117	Normal		35.36	76.36	56.43
118	Normal		30.29	178.50	21.07
119	Prostate hypertrophy		98.91	630.80	40.11
120	Normal		40.18	177.30	67.20
121	SpaceOAR		109.66	143.22	57.21
122	SpaceOAR		43.47	204.03	29.81
123	Prostate hypertrophy		53.08	98.81	136.07
124	Normal		27.20	281.91	34.15
125	Normal		41.13	375.57	48.68
126	Normal		50.91	216.65	45.31
127	Prostate hypertrophy		153.81	174.37	73.21
128	Morbid obesity		88.43	180.12	51.52
129	Normal		42.12	113.28	67.47
130	SpaceOAR		52.58	130.23	109.78
131	Foley catheter	Prostate hypertrophy	145.63	140.90	24.60

Supplementary Table 3. Detailed list of structure comparisons per individual for prostate, rectum, bladder, and bilateral femoral heads (femur_L, femur_R). Comparisons all versus expert clinician-delineated contours via Dice similarity coefficient (DSC). Results are for model-based (MB), deep learning (DL), and atlas-based (AB) algorithms.

ID	Prostate			Rectum			Bladder			Femur_L			Femur_R		
	MB	DL	AB	MB	DL	AB	MB	DL	AB	MB	DL	AB	MB	DL	AB
1	0.07	0.86	0.57	0.10	0.55	0.52	0.11	0.95	0.95	0.10	0.92	0.96	0.32	0.90	0.95
2	0.03	0.82	0.71	0.16	0.47	NA	0.22	0.93	0.64	0.89	0.95	0.93	0.94	0.93	0.95
3	0.63	0.81	0.52	0.40	0.47	0.41	0.48	0.88	0.58	0.91	0.96	0.95	0.95	0.92	0.96
4	0.86	0.91	0.41	0.47	0.66	0.68	0.94	0.96	0.85	0.93	0.92	0.94	0.96	0.90	0.95
5	0.64	0.72	0.00	0.48	0.63	0.42	0.85	0.93	NA	0.91	0.85	0.27	0.90	0.84	0.04
6	0.47	0.51	0.47	0.31	0.33	0.31	0.52	0.91	0.47	0.89	0.95	0.94	0.93	0.92	0.92
7	0.46	0.82	NA	0.65	0.68	0.52	0.27	0.91	NA	0.95	0.90	0.78	0.96	0.87	0.28
8	0.65	0.73	0.38	0.28	0.66	0.51	0.91	0.86	0.84	0.93	0.91	0.93	0.94	0.91	0.93
9	0.75	0.84	0.47	0.17	0.38	0.37	0.96	0.98	0.87	0.82	0.72	0.79	0.83	0.73	0.82
10	NA	NA	0.42	NA	NA	0.31	NA	NA	0.24	NA	NA	0.94	NA	NA	NA
11	0.76	0.84	0.44	0.56	0.62	0.52	0.93	0.92	0.73	0.91	0.91	0.94	0.94	0.89	0.93
12	0.21	0.61	0.25	0.40	0.52	0.43	0.12	0.66	NA	0.87	0.78	0.80	NA	NA	NA
13	0.68	0.82	0.00	0.44	0.55	0.11	0.32	0.86	NA	0.94	0.91	0.00	0.91	0.86	0.00
14	0.28	0.24	0.34	0.19	0.22	0.54	0.05	0.05	0.90	0.50	0.47	0.92	0.57	0.54	0.94
15	0.68	0.84	0.76	0.50	0.70	0.52	0.95	0.95	0.71	0.90	0.92	0.53	0.96	0.89	0.81
16	0.75	0.83	0.00	0.41	0.56	0.31	0.82	0.92	NA	0.96	0.93	0.54	0.91	0.88	0.38
17	0.70	0.74	0.55	0.17	0.31	0.33	0.92	0.92	0.56	0.93	0.96	0.94	0.97	0.89	0.95
18	0.71	0.76	0.18	0.41	0.73	NA	0.94	0.94	0.62	0.93	0.91	0.91	0.89	0.87	0.90
19	0.79	0.80	0.46	0.36	0.62	0.45	0.90	0.94	0.84	0.95	0.89	0.90	0.96	0.89	0.95
20	0.69	0.88	0.58	0.45	0.57	0.48	0.56	0.95	0.87	0.56	0.70	0.45	0.91	0.85	0.91
21	0.76	0.82	0.58	0.13	0.56	0.47	0.95	0.91	0.90	0.86	0.95	0.92	0.90	0.95	0.92
22	0.57	0.70	0.00	0.36	0.64	0.00	0.25	0.57	NA	0.91	0.93	0.00	0.91	0.94	NA
23	0.82	0.73	NA	0.25	0.64	0.00	0.46	0.78	NA	0.92	0.95	0.38	0.95	0.93	0.20
24	0.81	0.84	0.00	0.39	0.40	0.00	0.83	0.90	NA	0.90	0.94	0.00	0.92	0.96	0.00
25	0.52	0.83	NA	0.44	0.51	NA	0.74	0.92	NA	0.84	0.74	0.82	0.79	0.75	0.50
26	0.32	NA	0.00	0.29	NA	0.18	0.34	NA	NA	0.94	NA	0.00	NA	NA	NA
27	0.33	0.68	0.44	0.33	0.59	0.41	0.00	0.86	0.47	0.94	0.89	0.93	0.94	0.87	0.93
28	0.09	0.00	0.25	0.28	0.54	0.32	0.70	0.56	0.63	0.34	0.47	0.30	0.20	0.09	0.14
29	0.59	0.56	0.51	0.75	0.79	0.63	0.91	0.92	0.81	0.87	0.95	0.92	0.94	0.94	0.94
30	0.58	0.67	0.31	0.56	0.55	0.35	0.90	0.96	0.91	0.88	0.94	0.93	0.91	0.95	0.94
31	0.50	0.76	0.46	0.39	0.53	0.50	0.96	0.95	0.92	0.95	0.87	0.92	0.94	0.86	0.93
32	0.38	0.88	0.73	0.37	0.43	0.40	0.06	0.97	0.79	0.84	0.93	0.90	0.86	0.96	0.86
33	0.63	0.85	0.00	0.32	0.52	NA	0.58	0.94	NA	0.92	0.90	0.00	0.95	0.86	NA
34	0.46	0.76	0.63	0.25	0.32	0.35	0.96	0.94	0.71	0.77	0.70	0.73	0.78	0.69	0.77
35	0.17	0.67	0.23	0.40	0.32	0.30	0.01	0.87	NA	0.09	0.27	0.23	0.96	0.89	0.93
36	0.65	0.35	0.24	0.45	0.62	0.57	0.59	0.40	0.46	0.93	0.83	0.88	0.89	0.81	0.87
37	0.02	0.73	0.00	0.20	0.56	0.43	0.96	0.95	0.08	0.88	0.93	0.88	0.91	0.94	0.71
38	0.53	0.85	0.65	0.51	0.67	0.33	0.43	0.88	0.43	0.90	0.90	0.77	0.95	0.90	0.80
39	0.73	0.83	0.54	0.40	0.53	0.41	0.94	0.89	0.93	0.96	0.94	0.92	0.94	0.87	0.93
40	0.50	0.53	0.00	0.08	0.35	NA	0.91	0.70	0.00	0.84	0.75	0.00	0.81	0.77	0.00
41	0.08	0.79	0.00	0.34	0.50	NA	0.29	0.87	0.00	0.92	0.93	0.00	0.96	0.90	0.00
42	0.39	0.81	0.75	0.62	0.38	0.29	0.54	0.84	0.45	0.96	0.88	0.89	0.91	0.86	0.88
43	0.00	0.00	0.53	0.00	0.00	NA	0.00	0.00	0.48	0.00	0.00	0.90	0.00	0.00	0.89
44	0.09	0.87	0.65	0.10	0.37	0.26	0.75	0.95	0.93	0.87	0.91	0.94	0.95	0.89	0.95
45	0.17	0.85	0.53	0.39	0.44	0.42	0.67	0.95	0.81	0.93	0.89	0.95	0.95	0.88	0.96
46	0.41	0.72	0.51	0.10	0.23	0.29	0.84	0.88	0.75	0.80	0.73	0.80	0.71	0.66	0.75
47	0.47	0.84	0.59	0.59	0.68	0.62	0.94	0.94	0.94	0.96	0.88	0.90	0.95	0.87	0.93
48	0.51	0.79	0.45	0.16	0.46	0.38	0.53	0.95	0.69	0.91	0.90	0.94	0.94	0.90	0.94
49	0.11	0.00	0.21	0.12	0.28	0.25	0.37	0.41	0.64	0.08	0.11	0.10	0.17	0.08	0.11

Supplementary Table 4. Detailed list of structure comparisons per individual for prostate, rectum, bladder, and bilateral femoral heads (femur_L, femur_R). Comparisons all versus clinician-delineated contours via Mean surface distance (MSD) in millimeters. Results are for model-based (MB), deep learning (DL), and atlas-based (AB) algorithms.

ID	Prostate			Rectum			Bladder			Femur_L			Femur_R		
	MB	DL	AB	MB	DL	AB	MB	DL	AB	MB	DL	AB	MB	DL	AB
1	21.44	2.34	6.01	18.94	8.04	11.30	22.83	1.40	1.46	19.63	1.42	0.83	15.40	1.67	1.00
2	18.83	3.01	4.03	13.94	8.47	NA	14.70	2.02	16.04	2.32	1.07	1.34	1.37	1.32	1.17
3	5.50	2.52	5.09	7.31	10.81	15.65	6.99	2.51	13.87	1.82	0.96	0.93	1.28	1.48	0.85
4	3.34	2.20	9.40	9.37	7.14	7.90	1.89	1.57	5.06	1.51	1.72	1.59	1.27	2.02	1.28
5	5.63	4.15	62.33	11.53	6.57	9.57	2.63	1.82	NA	1.55	3.06	9.73	1.61	3.83	19.72
6	11.71	8.84	9.31	9.66	14.95	21.76	11.77	2.27	9.63	2.50	0.95	0.90	1.46	1.38	1.22
7	14.14	4.13	NA	4.30	4.82	12.86	13.72	2.20	NA	1.27	2.18	5.07	1.09	3.76	11.48
8	6.37	4.48	9.42	15.43	5.88	9.80	1.89	3.01	3.60	1.48	1.85	1.46	1.27	1.96	1.55
9	3.75	2.57	6.27	14.66	11.93	13.58	1.58	1.12	3.28	2.79	9.47	6.31	2.85	9.27	4.87
10	NA	NA	5.65	NA	NA	17.55	NA	NA	24.81	NA	NA	1.09	NA	NA	NA
11	4.58	3.00	8.06	6.08	6.17	8.74	1.52	1.71	6.76	1.49	1.37	1.05	1.12	1.71	1.20
12	16.11	4.82	9.59	12.26	14.27	12.11	28.86	11.74	NA	2.17	5.17	6.50	NA	NA	NA
13	4.78	2.90	124.63	9.52	10.31	38.95	17.50	3.06	NA	1.27	1.56	65.08	1.56	2.93	71.73
14	18.56	20.81	12.65	12.98	12.73	6.03	28.74	28.93	3.54	8.59	9.35	1.55	7.78	9.08	1.24
15	7.06	3.52	3.83	9.22	7.35	12.67	1.75	1.67	22.10	2.07	1.56	9.81	1.03	1.88	2.88
16	3.50	2.59	121.34	9.81	11.96	15.60	3.26	2.02	NA	1.24	1.32	14.67	1.62	2.21	10.00
17	4.58	3.42	5.10	22.05	17.49	17.98	1.89	1.96	21.37	1.41	1.05	1.15	1.07	2.27	0.93
18	4.47	3.79	10.55	12.39	5.57	NA	1.92	1.96	6.17	1.62	1.59	1.78	1.84	2.71	2.10
19	3.76	3.52	8.09	8.83	4.88	13.80	2.28	1.77	3.98	1.18	1.86	1.84	1.01	1.98	1.06
20	5.52	2.23	5.53	12.06	9.82	16.20	10.29	1.75	3.35	19.42	10.80	12.09	1.62	3.47	1.84
21	4.05	3.26	5.97	23.23	8.79	6.97	1.40	2.01	3.56	4.44	1.13	1.24	2.78	1.14	1.21
22	7.07	5.77	125.06	12.03	5.98	101.68	19.25	6.34	NA	1.86	1.49	97.61	1.83	1.52	NA
23	3.63	4.50	NA	17.17	6.48	89.74	12.58	4.59	NA	1.63	1.22	8.98	1.24	1.27	19.40
24	2.67	2.31	129.92	12.50	13.22	98.24	2.91	1.85	NA	2.41	1.25	48.50	1.78	1.24	62.42
25	8.39	3.10	NA	9.57	14.96	NA	4.79	1.80	NA	2.29	6.89	3.33	3.51	5.77	8.87
26	13.38	NA	125.54	13.98	NA	31.64	12.58	NA	NA	1.26	NA	81.29	NA	NA	NA
27	10.74	3.85	6.13	14.20	10.74	21.68	25.99	2.92	16.31	1.32	2.23	1.44	1.19	2.65	1.42
28	19.94	30.00	11.23	14.65	9.71	13.75	6.44	9.48	4.98	22.17	13.28	16.76	46.41	20.87	33.98
29	9.47	7.62	8.24	3.99	3.77	6.19	2.41	2.26	6.38	2.74	1.32	1.68	1.51	1.46	1.40
30	6.32	5.35	10.17	7.12	6.85	12.98	2.72	1.35	2.55	2.48	1.25	1.20	1.65	1.31	1.12
31	9.73	3.88	8.66	11.95	9.59	12.56	1.53	1.56	2.64	1.30	2.74	1.72	1.37	3.49	1.47
32	11.92	2.34	3.90	10.21	12.81	14.81	40.74	1.32	4.63	5.05	1.52	1.63	4.28	1.01	2.30
33	5.24	2.13	155.40	9.99	8.57	NA	9.63	1.80	NA	1.59	1.90	182.68	1.36	3.19	NA
34	10.08	3.71	5.14	19.47	17.28	20.76	1.33	1.60	13.03	3.75	8.43	9.06	3.73	10.10	6.98
35	21.38	4.71	9.18	10.07	16.06	14.58	33.94	2.74	NA	12.58	12.61	13.13	1.20	2.34	1.56
36	7.61	10.59	15.00	10.19	7.16	9.70	6.73	13.54	15.27	1.25	4.06	2.75	1.79	5.05	3.50
37	26.08	4.56	50.09	22.86	14.06	18.82	1.57	1.69	23.42	2.50	1.34	1.75	1.84	1.22	4.27
38	5.99	2.16	5.29	7.72	5.56	17.91	11.12	2.91	12.56	1.78	1.51	3.93	1.15	1.67	3.67
39	3.56	2.48	5.13	12.09	9.77	15.45	1.38	2.10	1.43	1.18	1.27	1.75	1.24	2.97	1.44
40	11.15	5.74	120.90	20.91	13.28	NA	1.73	5.68	115.96	2.30	6.05	61.26	3.32	5.17	54.08
41	25.52	2.82	141.90	18.22	13.93	NA	21.89	2.60	139.11	1.44	1.32	51.76	1.09	1.99	55.34
42	9.28	2.74	2.97	5.03	13.79	25.84	11.52	3.13	17.80	1.12	2.34	2.62	1.41	2.48	2.73
43	7.21	4.15	4.74	18.28	10.47	NA	6.19	2.71	19.28	4.01	1.20	1.59	1.86	1.13	1.82
44	20.87	2.15	4.57	17.76	11.41	15.65	7.06	1.66	2.55	2.61	1.54	1.05	1.09	1.85	1.03
45	17.32	2.62	6.89	7.72	11.57	12.33	7.96	1.41	6.36	1.37	1.93	1.06	1.08	2.20	0.99
46	6.46	2.83	6.64	23.27	16.99	22.37	2.74	2.46	8.00	2.31	6.15	5.54	4.16	8.46	7.31
47	8.16	2.59	4.95	7.85	8.29	12.50	1.64	1.97	2.29	1.13	2.25	1.96	1.21	2.96	1.56
48	8.47	3.23	7.36	22.03	13.20	27.41	10.35	1.84	11						

Supplementary Table 5. Detailed list of structure comparisons per individual for prostate, rectum, bladder, and bilateral femoral heads (femur_L, femur_R). Comparisons all versus clinician-delineated contours via 95% Hausdorff distance (HD) in millimeters. Results are for model-based (MB), deep learning (DL), and atlas-based (AB) algorithms.

ID	Prostate			Rectum			Bladder			Femur_L			Femur_R		
	MB	DL	AB	MB	DL	AB	MB	DL	AB	MB	DL	AB	MB	DL	AB
1	38.81	5.21	10.68	58.16	34.08	40.22	44.79	3.10	4.73	38.03	3.12	2.29	37.12	4.02	3.38
2	33.01	6.58	9.21	46.98	30.56	NA	28.32	3.96	51.45	12.02	2.69	4.03	3.26	3.17	3.85
3	14.60	6.50	11.00	19.54	42.02	48.09	16.65	6.78	40.12	8.01	2.20	2.11	2.49	4.08	1.92
4	7.94	5.83	24.50	34.15	30.15	28.03	5.26	3.29	18.67	3.43	4.24	4.33	2.54	8.06	2.72
5	14.53	8.34	80.83	43.77	24.46	28.32	7.62	4.14	NA	4.28	20.19	28.06	4.53	26.14	40.26
6	31.82	24.49	25.01	32.00	55.66	58.56	44.65	7.05	24.53	16.00	2.31	2.71	6.08	3.05	3.89
7	35.76	11.03	NA	16.95	21.88	42.01	38.93	6.80	NA	2.62	12.03	24.00	2.26	26.31	35.77
8	14.27	11.22	23.18	51.18	23.88	31.93	4.92	10.00	12.52	4.08	7.07	4.15	2.49	8.19	4.42
9	8.66	6.00	15.39	40.15	37.33	40.58	2.80	2.29	15.73	16.99	51.59	31.18	18.26	51.89	26.36
10	NA	NA	15.62	NA	NA	48.65	NA	NA	60.14	NA	NA	3.11	NA	NA	NA
11	13.45	9.83	20.72	25.76	32.10	28.13	3.97	3.95	22.94	4.39	2.95	2.36	2.18	4.76	2.91
12	35.87	10.19	21.62	46.97	39.36	36.60	63.76	45.78	NA	10.06	31.62	36.51	NA	NA	NA
13	14.40	6.48	144.31	35.70	38.16	87.25	54.19	9.05	NA	2.51	4.08	94.95	3.46	16.00	108.07
14	42.84	48.08	29.87	43.91	34.76	16.84	53.37	53.55	16.13	19.60	21.23	6.20	18.06	24.25	3.02
15	22.35	13.63	12.34	36.20	35.18	44.07	5.47	3.14	85.98	11.18	6.05	38.24	2.23	8.17	12.07
16	8.81	5.49	142.23	25.17	38.28	50.16	12.66	4.99	NA	2.50	2.77	50.35	3.55	10.26	34.38
17	12.75	8.65	11.37	66.58	53.12	52.46	6.09	4.12	59.20	4.32	2.50	2.90	2.27	12.11	2.14
18	11.67	8.04	24.91	43.70	25.44	NA	3.85	4.86	22.40	4.22	3.87	6.18	5.42	15.02	6.08
19	10.23	8.86	18.52	36.00	24.26	43.89	6.99	3.66	14.03	2.38	7.03	8.61	2.21	8.06	2.52
20	16.70	5.68	15.15	45.57	36.18	54.32	29.47	3.69	10.01	96.01	63.53	55.70	4.58	21.41	9.25
21	11.13	8.58	11.79	71.14	34.07	18.36	3.14	4.31	16.05	28.44	2.53	2.75	18.11	2.47	2.54
22	17.00	21.25	150.01	38.29	24.25	132.56	54.16	20.52	NA	5.07	3.59	144.87	4.95	3.65	NA
23	11.28	14.98	NA	57.14	26.18	114.20	46.83	22.22	NA	6.08	2.88	26.27	2.87	2.68	50.68
24	6.58	5.61	148.35	37.28	40.63	118.14	10.01	4.99	NA	13.76	3.05	85.99	6.15	2.81	104.29
25	23.83	6.78	NA	31.95	66.44	NA	24.93	4.03	NA	13.00	42.37	11.44	24.04	34.83	26.79
26	22.37	NA	152.00	39.55	NA	70.40	33.53	NA	NA	2.44	NA	112.40	NA	NA	NA
27	26.05	8.39	12.33	48.01	46.40	66.03	58.39	6.24	34.57	2.82	12.06	4.35	2.33	16.24	4.47
28	38.16	46.23	27.18	55.82	36.95	38.06	25.42	31.22	21.34	73.30	50.04	58.85	119.41	66.48	108.29
29	28.85	21.74	19.71	17.90	18.01	21.01	7.49	6.09	27.18	12.78	3.10	4.70	3.54	3.34	3.66
30	14.60	14.76	21.68	34.05	29.55	44.40	10.83	2.70	8.77	14.01	2.84	2.56	5.26	2.89	2.62
31	22.35	8.66	15.70	38.06	36.14	47.52	3.01	3.11	12.06	2.46	16.35	7.95	3.01	23.15	4.01
32	30.62	5.85	9.39	30.44	41.51	47.31	86.07	2.65	17.33	32.57	4.76	4.46	28.50	2.29	10.08
33	13.33	4.24	176.53	29.09	35.24	NA	28.79	3.62	NA	4.06	6.13	224.35	2.66	18.10	NA
34	24.23	9.73	13.68	58.79	52.54	60.91	3.52	3.81	47.81	22.04	45.82	40.32	23.31	56.02	34.00
35	51.75	15.92	27.36	38.56	40.22	38.59	75.70	6.08	NA	33.34	49.95	36.97	2.30	12.67	4.41
36	25.03	39.07	40.03	38.41	30.00	34.52	29.76	48.64	44.09	2.59	27.35	16.95	6.38	33.26	22.16
37	44.74	10.28	70.66	75.89	47.77	53.66	3.52	3.09	58.87	13.01	2.90	6.41	7.01	2.77	18.41
38	22.27	4.96	10.40	18.23	23.13	57.35	32.75	8.53	34.25	6.20	3.43	15.82	2.29	3.22	14.59
39	8.39	6.50	10.24	46.68	37.37	50.54	2.81	7.83	3.50	2.29	2.70	8.44	2.41	19.01	4.42
40	28.78	16.05	143.73	51.24	44.16	NA	5.54	22.92	134.12	12.99	35.09	111.82	22.18	30.17	97.13
41	62.41	8.45	160.08	64.87	52.91	NA	61.47	5.40	154.70	3.33	2.97	85.12	2.23	8.42	92.56
42	27.07	7.93	6.11	18.52	46.50	66.96	49.12	9.39	46.16	2.31	12.21	15.58	3.12	12.23	16.03
43	19.53	14.40	11.79	42.39	34.76	NA	18.20	8.01	58.56	24.22	2.66	4.54	6.24	2.25	4.81
44	41.96	6.38	8.53	46.03	42.54	51.05	21.86	3.20	9.18	14.29	3.67	2.56	2.51	6.21	2.72
45	34.50	7.41	12.53	32.35	40.99	41.83	30.56	3.06	29.39	2.98	7.55	2.43	2.20	12.01	2.28
46	17.99	5.63	14.13	58.28	45.76	54.87	11.42	7.39	28.31	10.10	36.13	28.36	26.01	44.97	34.27