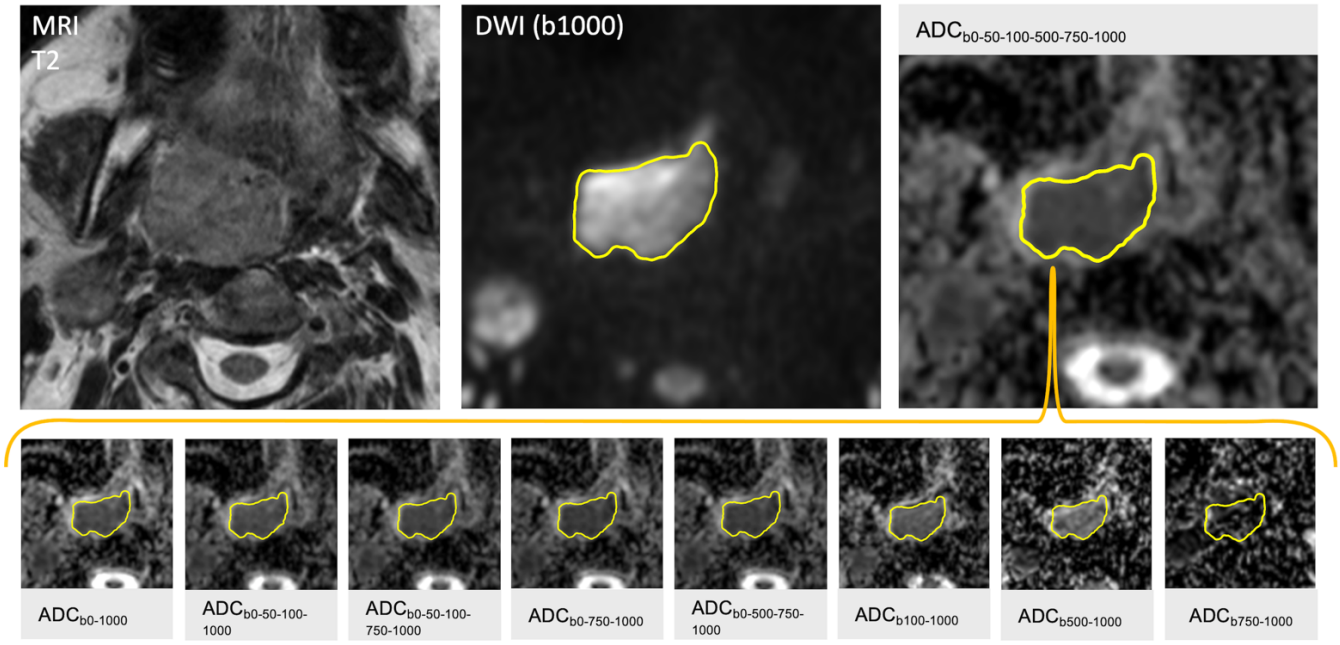
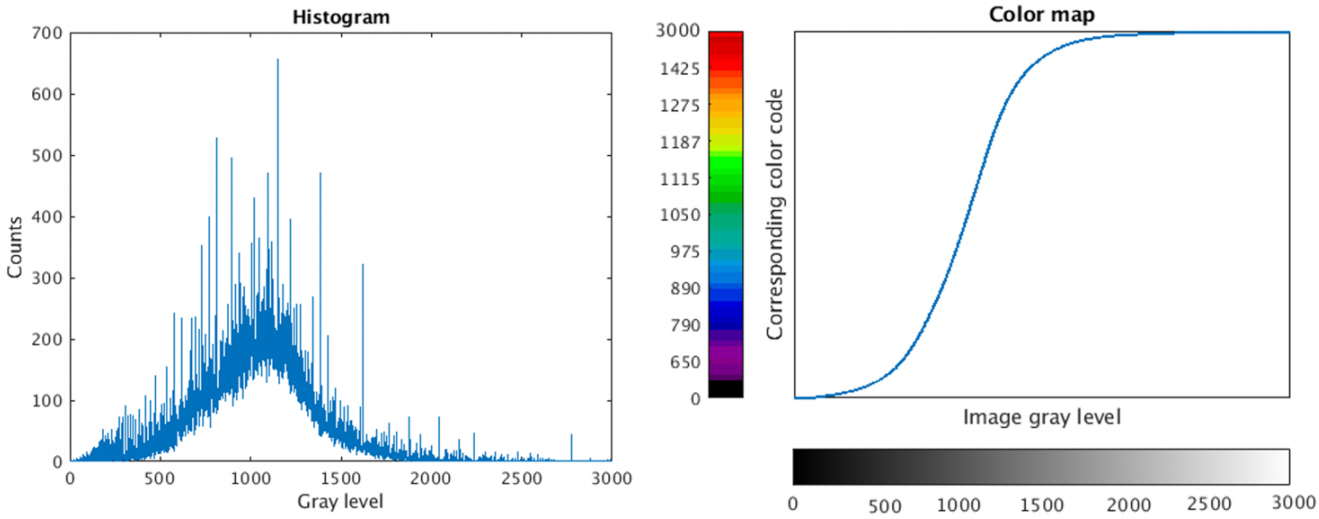
ONLINE SUPPLEMENTAL DATA

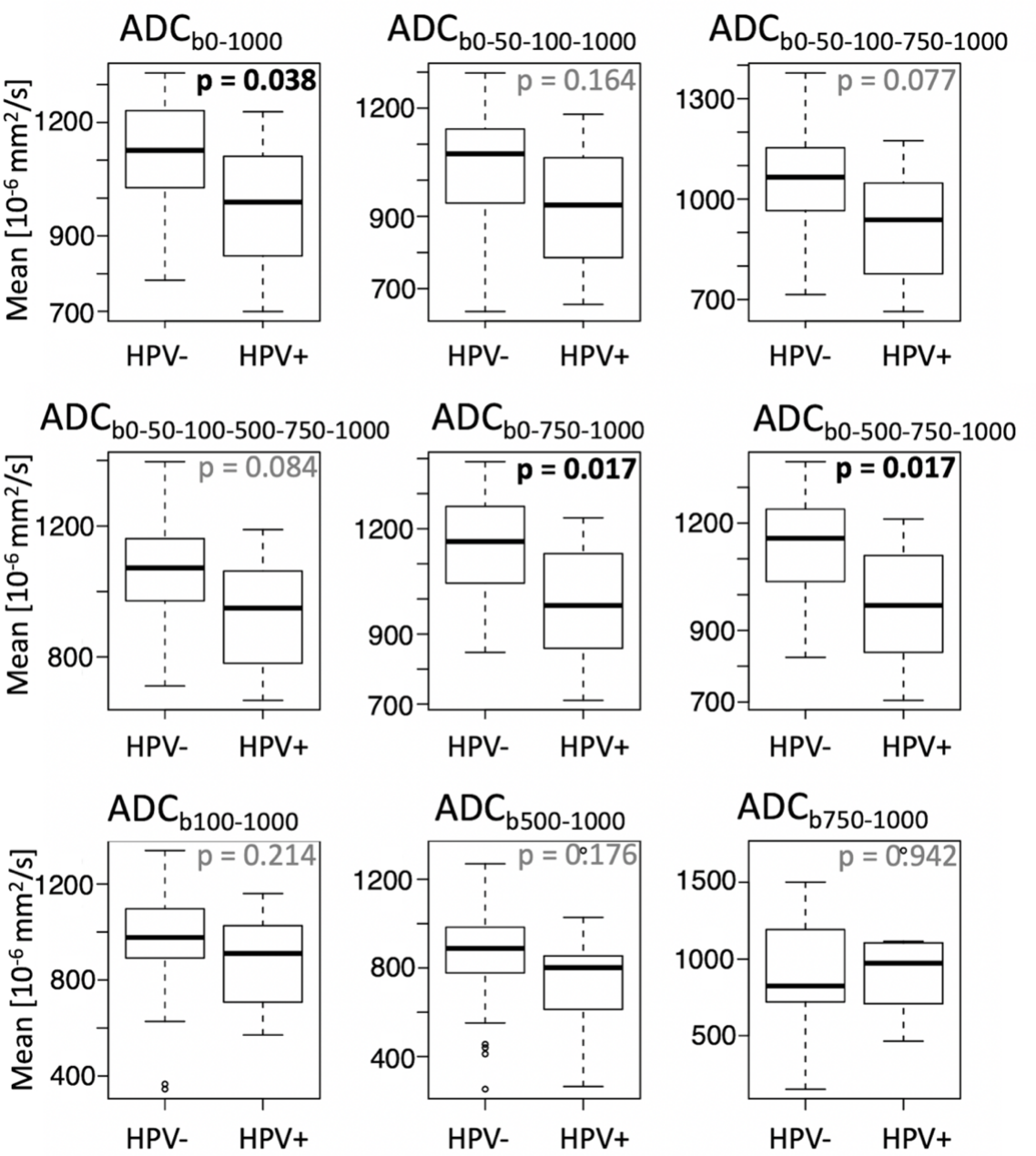
Supplemental Online Figure 1. Illustration of freehand contouring of an HPV+OPSCC across its largest cross-sectional area on T2, b=1000 and ADC images (upper figure row). The ADC image used for segmentation corresponded to the ADC map used in clinical routine generated automatically by the MRI scanner software (6 b-values). The same ROI was then copied on all studied ADC maps obtained with different b-value combinations (lower row).



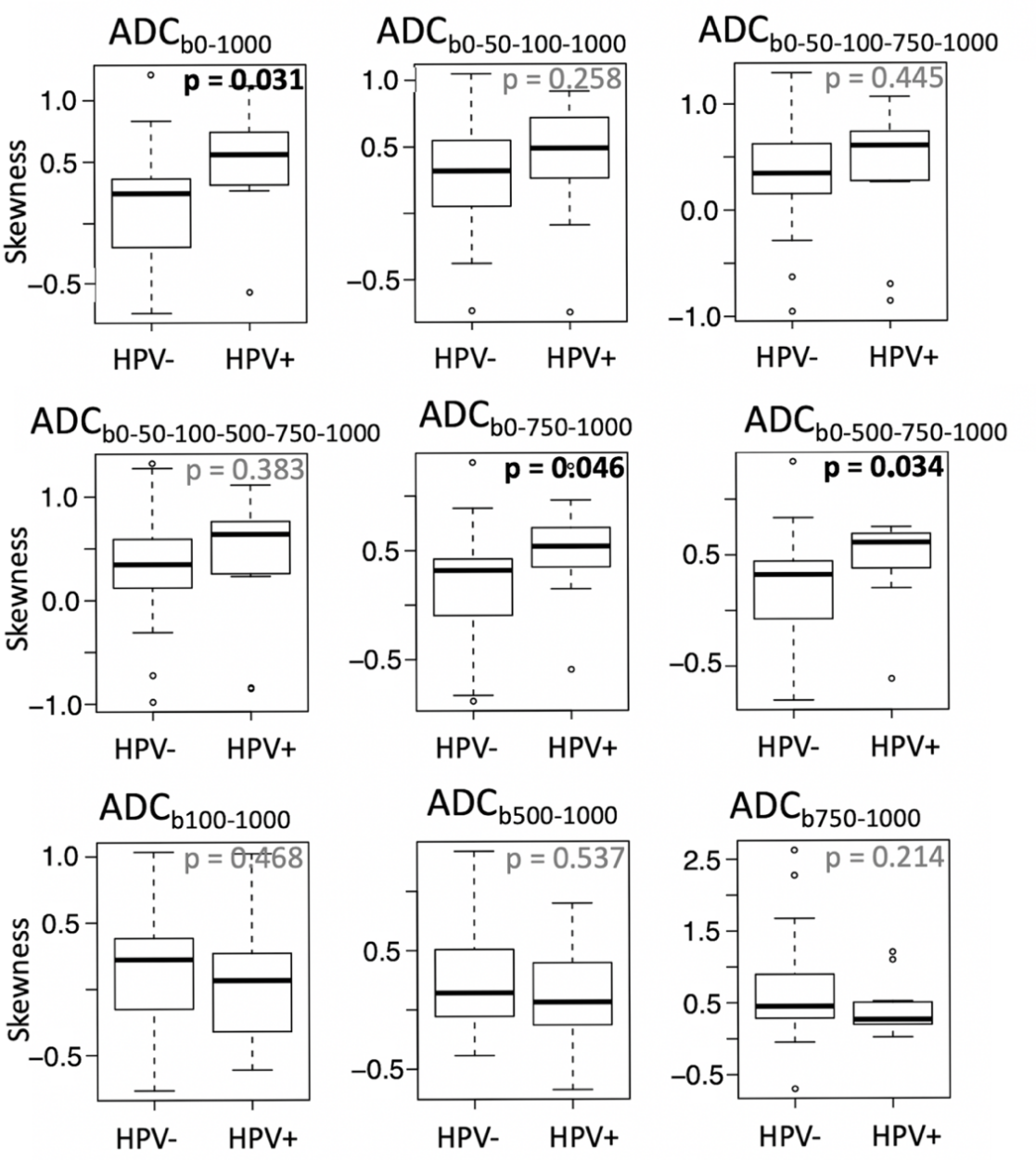
Supplemental Online Figure 2. To facilitate visual comparison between tumor ROIs on different ADC maps, a customized color-map for all ADC gray levels was used. A histogram was computed for all tumor ROIs of all patients. Then the normalized cumulative sum of this histogram was calculated, which resulted in a corresponding color-map for all gray levels. The UCLA-colormap was chosen to display the original gray levels (unit in x 10^{-6} mm^2/s). With this procedure, the range of gray levels containing the majority of pixels are displayed with a wide range of colors allowing improved visual differentiation between small changes in ADC pixel values, which may be more difficult to observe on standard gray scale images.



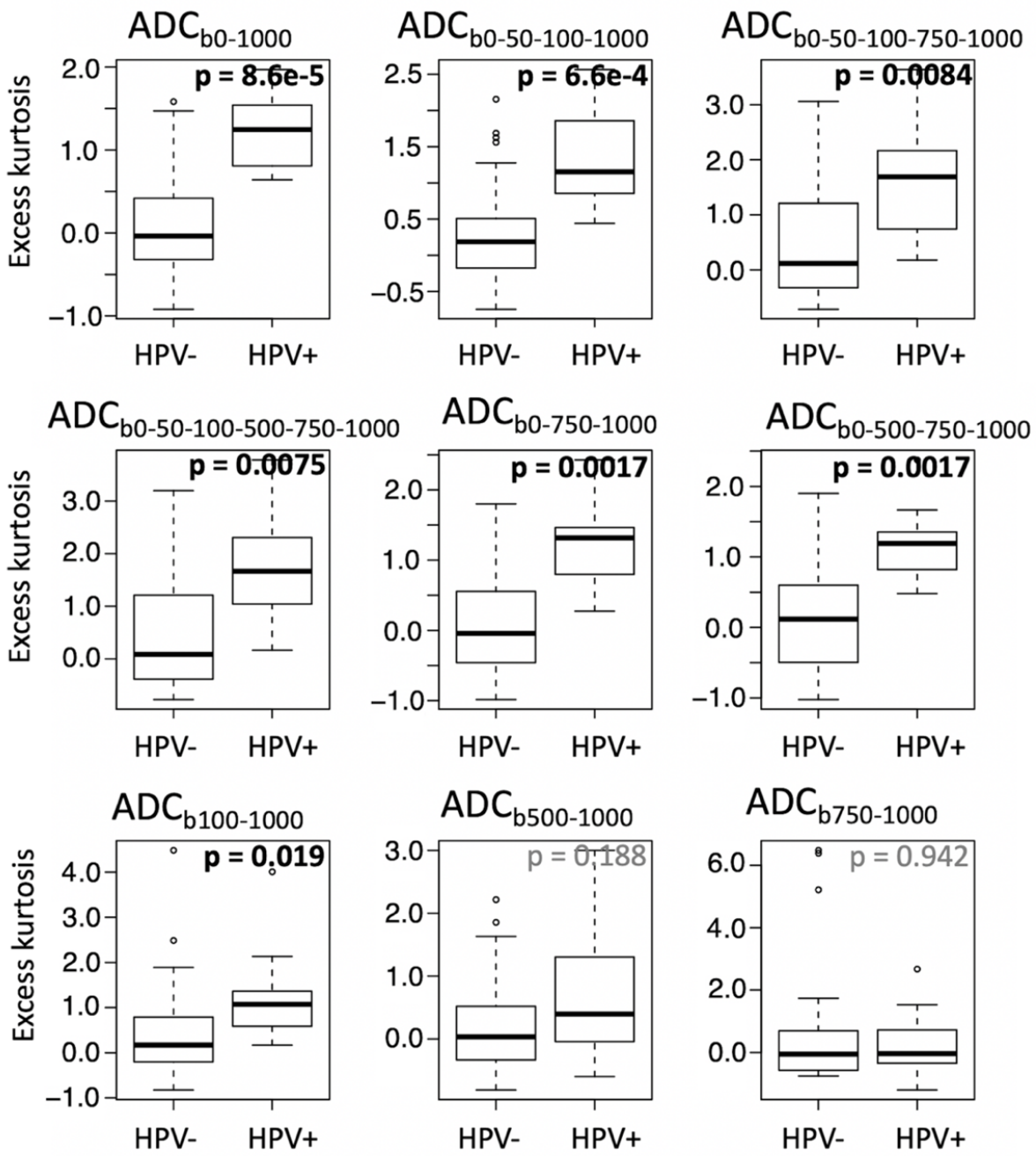
Supplemental Online Figure 3. Boxplots of ADC mean values in HPV- and HPV+OPSCC calculated for each combination of b-values. ADC mean values are indicated in 10^{-6} x mm^2/s. Median values are shown as bold lines, the 1st and 3rd quartiles are the limit of the box, whiskers extent to ±1.5 x interquartile range, and outliers are represented with circles. P values were calculated with the Mann-Whitney-Wilcoxon test. ADC mean values in HPV- versus HPV+OPSCC were significantly lower only for $ADC_{b0\text{-}1000}$, $ADC_{b0\text{-}750\text{-}1000}$ and $ADC_{b0\text{-}500\text{-}750\text{-}1000}$, however, not for all other six combinations. In particular, ADC mean values in HPV- and HPV+OPSCC were similar on ADC maps calculated with $b \geq 100 s/mm^2$.



Supplemental Online Figure 4. Boxplots of ADC skewness in HPV- and HPV+OPSCC calculated for each combination of b-values. Median values are shown as bold lines, the 1st and 3rd quartiles are the limit of the box, whiskers extent to ±1.5 x interquartile range, and outliers are represented with circles. P values were calculated with the Mann-Whitney-Wilcoxon test. ADC skewness in HPV-OPSCC was significantly lower than in HPV+OPSCC only for three combinations of b-values ($ADC_{b0\text{-}1000}$, $ADC_{b0\text{-}750\text{-}1000}$, $ADC_{b0\text{-}500\text{-}750\text{-}1000}$), however, not for all other combinations. In particular, ADC skewness in HPV- and HPV+OPSCC was similar on perfusion-insensitive ADC maps calculated with b≥100s/mm².



Supplemental Online Figure 5. Boxplots of ADC excess kurtosis in HPV- and HPV+OPSCC calculated for each combination of b-values. Median values are shown as bold lines, the 1st and 3rd quartiles are the limit of the box, whiskers extent to ±1.5 x interquartile range, and outliers are represented with circles. P values were calculated with the Mann-Whitney-Wilcoxon test. Note that ADC excess kurtosis in HPV- versus HPV+OPSCC was significantly lower ($p<0.05$) for all combinations of b-values with the exception of ADC excess kurtosis calculated with $b \geq 500 s/mm^2$ ($ADC_{b500-1000}$ and $ADC_{b750-1000}$).



Supplemental Online Table 1. Same data as in **Table 4** but with AUCs calculated with a leave-one-out cross-validation and with LDA as model. P values to distinguish HPV+ from HPV-OPSCC were calculated with the Mann-Whitney-Wilcoxon test. Sensitivity, specificity and accuracy were calculated with the optimal threshold extracted from LDA scores.

ADC map	Feature	AUC	p-value	TP	FP	TN	FN	Sensitivity	Specificity	Accuracy	Optimal threshold
$ADC_{b0\text{-}1000}$	ADC mean	0.719	0.021	8	7	16	3	0.727	0.696	0.706	1049.295[a]
	ADC skewness	0.731	0.016	8	6	17	3	0.727	0.739	0.735	0.348
	ADC kurtosis[b]	0.889	<0.001	11	4	19	0	1.000	0.826	0.882	0.581
$ADC_{b0\text{-}750\text{-}1000}$	ADC mean	0.743	0.012	8	7	16	3	0.727	0.695	0.706	1077.717[a]
	ADC skewness	0.715	0.023	8	5	18	3	0.727	0.783	0.765	0.459
	ADC kurtosis[b]	0.826	0.001	10	5	18	1	0.909	0.783	0.823	0.600
$ADC_{b0\text{-}500\text{-}750\text{-}1000}$	ADC mean	0.743	0.012	8	5	18	3	0.727	0.783	0.765	1032.304[a]
	ADC skewness	0.727	0.017	8	4	19	3	0.727	0.826	0.794	0.515
	ADC kurtosis[b]	0.826	0.001	10	5	18	1	0.909	0.783	0.823	0.607
$ADC_{b100\text{-}1000}$	ADC mean	0.632	0.114	8	8	15	3	0.727	0.652	0.676	952.723[a]
	ADC skewness	0.581	0.234	8	11	12	3	0.727	0.522	0.588	0.211
	ADC kurtosis[b]	0.751	0.009	9	7	16	2	0.818	0.696	0.735	0.442

a=ADC mean thresholds in $x10^{-6}mm^2/s$; b=excess kurtosis.