

Table I: Dataset Composition

Class	Description	Train	Val	Test	Total	%
Adenocarcinoma	Peripheral NSCLC	420	115	51	586	22.0%
Squamous Cell	Central, smoking	312	115	39	466	17.5%
Large Cell	Poorly diff.	224	115	28	367	13.8%
Benign	Non-malignant	144	115	18	277	10.4%
Normal	Healthy tissue	760	115	95	970	36.4%
Total		1860	575	231	2666	100%