

	Train	Test
Certificates	13,330	6,665
Lines	32,714	14,834
Tokens	90,442	42,819
Total ICD codes	39,334	18,928
Unique ICD codes	1,256	900
Unique unseen ICD codes	-	157