Head and Neck 5000 - Disease category and stage of disease

	TNM stage				
Diagnosis	**I**	**II**	**III**	**IV**	**Total cases**
Oral cavity	432 (33.7%)	290 (22.6%)	103 (8.0%)	458 (35.7%)	1,288
Oropharynx	83 (4.4%)	184 (9.8%)	268 (14.2%)	1,353 (71.7%)	1,896
Nasopharynx	8 (6.5%)	27 (22.0%)	46 (37.4%)	42 (34.2%)	124
Hypopharynx	14 (5.9%)	36 (15.2%)	34 (14.4%)	153 (64.6%)	237
Larynx	436 (41.0%)	272 (25.6%)	174 (16.4%)	182 (17.1%)	1,065
Thyroid	125 (48.8%)	32 (12.5%)	46 (18.0%)	53 (20.7%)	261
Nasal cavity	18 (29.5%)	16 (26.3%)	7 (11.5%)	20 (32.8%)	61
Sinuses	0 (0.0%)	2 (5.0%)	4 (10.0%)	34 (85.0%)	40
Salivary glands	71 (35.2%)	35 (17.3%)	25 (12.4%)	71 (35.2%)	208
Total	1,187 (23.0%)	894 (17.4%)	707 (13.7%)	2,366 (45.9%)	5,154

Where pathology forms were available, the ICD code from the pathology form was used. Where pathology forms were not available, the ICD code from the clinical baseline data capture form was used.

Categories comprised of the following ICD codes:

- Oral cavity = C00, C02, C03, C04, C05, C06 (excluding C02.4, C05.1, C05.2, C05.8)
- Oropharynx = C01, C02.4, C05.1, C05.2, C05.8, C09, C10 (excluding C10.1)
- Nasopharynx = C11
- Hypopharynx = C12, C13
- Larynx = C10.1, C32.0
- Thyroid = C73
- Nasal cavity = C30 (excluding C30.1)
- Sinuses = C31
- Salivary glands = C07, C08

The dataset contains data for two other categories of tumours of the head and neck:

- Primary of unknown origin = C80 (204 participants)
- Other = C14.0, C30.1, C41.1, C69.5 (20 participants)

Tumours which were identified as salivary glands based on histology but had a different ICD code were recoded as salivary gland tumours.

TNM 7th edition was used for all tumour sites. For HPV positive (based on HPV16E6 serology) oropharyngeal cancers, the International Collaboration on Oropharyngeal cancer Network for Staging (ICON-S) classification has also been used in the dataset. The numbers are as follows: stage I 608 (58.5%); stage II 220 (21.2%); stage III 209 (20.1%); stage IV 3 (0.3%).