medRxiv preprint doi: https://doi.org/10.1101/2024.12.05.24318538; this version posted December 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Incidence, prevalence, and survival of head and neck cancers in the United Kingdom from 2000-2021

Author names:

Andrea Miquel Dominguez[a]: MD amiquel.hj23.ics@gencat.cat

Eng Hooi Tan[b]:PhD cheryl.tan@ndorms.ox.ac.uk

Edward Burn[b]: PhD edward.burn@ndorms.ox.ac.uk

Antonella Delmestri[b]: PhD antonella.delmestri@ndorms.ox.ac.uk

Talita Duarte-Salles[c,d] : PhD tduarte@idiapjgol.org

Asieh Golozar[e,f]: PhD asieh.golozar@odysseusinc.com

Wai Yi Man[b]: MSc waiyi.man@ndorms.ox.ac.uk

Daniel Prieto-Alhambra[b,d]: PhD daniel.prietoalhambra@ndorms.ox.ac.uk

Francesc Xavier Avilés-Jurado[g]: MD faviles@clinic.cat

Danielle Newby[b]: PhD danielle.newby@ndorms.ox.ac.uk

Affiliations:

[a] Otorrinolaringology department, Hospital Joan XXIII de Tarragona, C/ Dr. Mallafrè Guasch, 4, 43005 Tarragona, Spain

[b] Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, OX3 7LD, United Kingdom

[c] Fundació Institut Universitari per a la recerca a l'Atenció Primària de Salut Jordi Gol i Gurina (IDIAPJGol), Barcelona, Spain

[d] Department of Medical Informatics, Erasmus University Medical Centre, 3015 GD Rotterdam, The Netherlands

[e] Odysseus Data Services, Cambridge, MA 02142, United States

[f] OHDSI Center at the Roux Institute, Northeastern University, Boston, MA 02115, USA

[g] Head Neck Tumors Unit. Hospital Clínic de Barcelona. Universitat de Barcelona. IDIBAPS. IISPV. Villarroel 170, 08036 Barcelona (Catalonia, Spain)

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2024.12.05.24318538; this version posted December 6, 2024. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

Corresponding author:

Professor Daniel Prieto Alhambra

Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, OX3 7LD, UK.

daniel.prietoalhambra@ndorms.ox.ac.uk

medRxiv preprint doi: https://doi.org/10.1101/2024.12.05.24318538; this version posted December 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

KEY POINTS

Question: What is the disease burden of the head and neck cancers (HNC) in UK from 2000-2021?

Findings: For all HNC combined, incidence and prevalence have increased with five-year survival slightly improving over time. Variation in results were observed for different subsites**.**

Meaning: Our findings show increases in HNC are likely due to ageing with increases in certain subsites due to other behavioural risk factors. The small increase in survival highlights more research is needed to improve earlier diagnosis leading to better patient outcomes.

medRxiv preprint doi: https://doi.org/10.1101/2024.12.05.24318538; this version posted December 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

ABSTRACT:

Importance: Incidence, prevalence, and survival are important measures to inform the management and provision of head and neck cancer care.

Objective To calculate the incidence, prevalence, and survival rates for head and neck cancers and subsites in the UK from 2000 to 2021.

Design, Setting, and Participants: This population-based cohort study uses routinely collected primary care data from the UK. Patients aged 18 years or older with at least 1 year of history registered in Clinical Practice Research Datalink (CPRD) GOLD or Aurum were included. Data were analyzed from January 2023 to March 2024.

Main Outcomes and Measures: Head and neck cancer incidence rates (IR), period prevalence (PP), and one-, five-, and ten-year survival after diagnosis between 2000 and 2021, stratified by age and calendar years.

Results: There were 12,455 HNC patients (male 69.2%) with a median age of 64 years. Crude incidence increased from 9.08 (7.88 to 10.42) per 100 000 person-years in 2000 to 15.59 (14.07 to 17.23) in 2021 in CPRD GOLD with similar rates in Aurum. Age standardization attenuated incidence rises for HNC subsites apart from oropharynx and tongue. Prevalence increased for both databases, from 0.04% in 2000 to 0.12% in 2019. HNC five-year survival increased from 53.8% (95% CI, 51.4% - 56.3%) in 2000-2004 to 58.7% (56.5 - 60.9%) in 2015 to 2019.

Conclusions and Relevance: HNC increases over recent decades are likely due to ageing with increases in specific subsites such as oropharyngeal cancers are due to other behavioural risk factors. Small improvements in survival highlights more research is needed to improve earlier diagnosis which will lead to better patient outcomes.

medRxiv preprint doi: https://doi.org/10.1101/2024.12.05.24318538; this version posted December 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

KEYWORDS:

Head and neck cancer, head and neck neoplasms, oral cancer, oncology, incidence, prevalence, cancer survival

ABBREVIATIONS:

IR; incidence rates, PP; period prevalence, HNC; head and neck cancers, KM; Kaplan-Meier, HPV, human papillomavirus

WORD COUNT: 2983

FIGURE & TABLES: 4

medRxiv preprint doi: https://doi.org/10.1101/2024.12.05.24318538; this version posted December 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

INTRODUCTION

Head and neck cancers (HNC) are a heterogenous group of cancers affecting the upper aerodigestive tract including the oral cavity, pharynx, larynx, paranasal sinuses, nasal cavity, and salivary glands. Combined, these cancers are the seventh most common cancer worldwide accounting for more than 660,000 new cases and 325,000 deaths each year [1.]

The etiology and epidemiology of HNC subsites can vary with distinct characteristics in terms of histology subtypes, symptoms, treatment approaches, and prognosis, making it necessary to consider them as separate entities [2]. The overall incidence of HNC has been shown to be increasing over time with marginal improvements in survival in the United Kingdom (UK) and in other countries[6,7,11] However, specific risk factors such as tobacco and alcohol consumption and human papillomavirus (HPV) infection have impacted the disease burden of the different HNC subtypes [9 10].

Understanding trends in incidence, prevalence, and overall survival of HNC and subsites is an important aspect to inform decisions regarding screening, prevention, treatment, and disease management. Due to changes in HPV infections and vaccinations, a decline in tobacco and alcohol usage in recent years, an up-to-date comprehensive assessment of the trends of HNC and subsites is lacking. The aim of this study is to describe HNC and subsite trends in terms of incidence, prevalence, and survival from 2000-2021 using primary care data from the UK.

METHODS

Study design, setting, and data sources.

We carried out a population cohort study using routinely collected primary care data from the UK. People with a diagnosis of HNC and a denominator population were identified from the CPRD GOLD database. We repeated the study using CPRD Aurum for comparison. Both databases contain pseudonymised patient-level information on demographics, lifestyle data, clinical diagnoses, prescriptions, and preventive care and are broadly representative[15]. Both databases were mapped to the Observational Medical Outcomes Partnership Common Data Model[16,17].

Study participants

medRxiv preprint doi: https://doi.org/10.1101/2024.12.05.24318538; this version posted December 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Eligible individuals were 18 years or older with one year of prior history, in the database from 1st January 2000. These individuals were followed up to whichever came first: HNC diagnosis, exit from database, date of death, or study end (December 31, 2021, for GOLD, and, due to data availability December 31, 2019, for Aurum). For survival analysis, individuals were followed from cancer diagnosis to either death, exit from database, or study end.

Outcome definitions

We used diagnostic codes to identify HNC and subsites. Diagnostic codes indicative of either nonmalignant cancer or metastasis were excluded (apart from prevalence analyses) as well as melanoma and lymphoma codes. Subsites of HNC were categorised as cancers of the oral cavity, tongue and lingual tonsils, nasal cavity and sinuses, salivary glands, hypopharynx, nasopharynx, oropharynx, and larynx. Clinical codelists are provided in supplement S1 with all analytical code in GitHub to enable reproducibility[58.] For survival analyses, mortality was defined as all-cause mortality from date of death.

Statistical methods

The characteristics of HNC patients were summarised, with median (IQR) for continuous variables and counts with percentages for categorical variables.

For incidence, number of events, observed time at risk, and incidence rates (IR) per 100,000 person years were summarised along with 95% confidence intervals (CI). Annualised IR were calculated as the number of incident cases as the numerator and number of person-years in the general population within that year as the denominator from 2000 to 2021. Age-standardized IR were calculated using the 2013 European Standard Population[(19)].

Period prevalence was calculated on 1st January for the years 2000 to 2021 with HNC patients as the numerator. The denominator was the number of patients in the respective years. The number of events and prevalence (%) were summarised along with 95% CI.

For survival analysis, we used the Kaplan-Meier method to estimate the overall survival probability with 95% CI. We estimated median survival and survival one, five, and ten years after diagnosis. Patients whose death and cancer diagnosis occurred on the same date were removed.

medRxiv preprint doi: https://doi.org/10.1101/2024.12.05.24318538; this version posted December 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

All results were stratified by sex and by age group. For survival analysis, we also stratified by calendar time of cancer diagnosis. To avoid possible patient re-identification, we do not report results with fewer than five cases. To replicate the results from GOLD, the same analysis was performed using Aurum, except for the stratification by calendar time of cancer diagnosis, which was conducted in GOLD only.

The statistical software R version 4.2.3 was used for analyses and the IncidencePrevalence [20] and survival R packages [41]. Data were analyzed from January 2023 to March 2024. Data were descriptive and no tests for statistical significance were performed.

RESULTS

Patient characteristics

Overall, there were 11,388,117 eligible patients, with at least one year of prior history from 2000-2021 in GOLD. Attrition tables are in the supplement S2. A summary of baseline patient characteristics of HNC patients is shown in Table 1. All study results for this study can be found in an interactive web application[42].

Table 1: Demographics of HNC patients at the time of diagnosis for CPRD GOLD.

medRxiv preprint doi: https://doi.org/10.1101/2024.12.05.24318538; this version posted December 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Database	**CPRD GOLD**
Number of patients	12455
Sex: Male (N [%])	8614 (69.2%)
Age (years) (Median [IQR])	64 (56 to 73)
Age Groups N (years) (%)	
18-29	95 (0.8%)
30-39	258 (2.1%)
40-49	1071 (8.6%)
50-59	3001 (24.1%)
60-69	3730 (29.9%)
70-79	2790 (22.4%)
80-89	1293 (10.4%)
90+	217 (1.7%)
Prior history, days	
median [p25 - p75]	3368 (1,818 to 5,119)
Smoking Status (5 years prior)	
Non-smoker	3187 (25.6%)
Former smoker	92 (0.7%)
Current smoker	5322 (42.7%)
Missing	3854 (30.9%)
General conditions (any time prior)	
Cerebrovascular disease	727 (5.8%)
Chronic liver disease	178 (1.4%)
Chronic obstructive lung disease	1217 (9.8%)
Coronary arteriosclerosis	115 (0.9%)
Dementia	163 (1.3%)
Depressive disorder	1655 (13.3%)
Diabetes mellitus	1097 (8.8%)
Gastroesophageal reflux disease	330 (2.6%)
Gastrointestinal haemorrhage	797 (6.4%)
Heart disease	1865 (15.0%)
Hyperlipidemia	996 (8.0%)
Hypertensive disorder	2966 (23.8%)
Ischemic heart disease	1029 (8.3%)
Osteoarthritis	1938 (15.6%)
Peripheral vascular disease	379 (3.0%)
Renal impairment	1067 (8.6%)
Venous thrombosis	451 (3.6%)

There were 12,455 HNC patients (69.2% males), with a median age of 64 (56 to 73) years. The highest percentage of patients were 60-69 years olds contributing to 30% of diagnosed patients with similar results in Aurum (Supplement S3).

medRxiv preprint doi: https://doi.org/10.1101/2024.12.05.24318538; this version posted December 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Characteristics by HNC subsite (Supplement S4) showed larynx (22%), tongue (21%), or oropharynx (18.6%) were most common with the rarest being nasal cavity and sinus cancers (2.1%). Patients were more likely to be male (69%), with laryngeal cancer having the highest proportion of males (81%). The highest proportion of diagnoses were those aged 60-69 years of age across all subsites apart from nasal cavity and sinuses and salivary gland where the highest proportion of patients were 70-79 years old. For nasopharyngeal cancer the highest proportion of patients were 50-59 years of age. Patients with cancers of the hypopharynx, larynx and oral cavity had higher percentages of comorbidities. For most subsites (larynx, oral cavity, tongue, oropharynx, hypopharynx), smokers made up the highest proportion of patients (37.3% to 53.7%).

Incidence rates stratified by calendar year, age, and sex.

In GOLD, crude overall IR of HNC from 2000 to 2021 was 14.2 (14.0 to 14.5) per 100,000 person-years. Males had higher overall IR (20.0 [19.6 to 20.4]) compared to females (8.6 [8.4 to 8.9]), with similar values in Aurum. The HNC subsite with the highest overall IR was laryngeal cancer (3.2 per 100,000 person-years) whereas nasal cavity and sinuses (0.3 per 100,000 for both databases) had the lowest. Across all subsites males had higher overall IR compared to females (supplement S5).

Crude annualised IR for HNC increased across the study period (Figure 1). For HNC subsites, cancers of the nasopharynx, oropharynx, salivary glands, and tongue increased over the study period. IR for hypopharyngeal cancer increased from 2000 to 2009 before stabilising in GOLD or reducing in Aurum. For cancers of the larynx and nasal cavity and sinuses, IR remained stable in GOLD with a decrease for laryngeal cancer in Aurum. For cancer of the oral cavity, IR gradually increased from 2000 to 2010-11 before dropping in 2012 before increasing. Stratification by sex revealed similar trends over time with higher IR for subsites in males apart from nasal cavity and sinuses cancers (Supplement S6). Age standardized IR show trends have attenuated or stabilised apart from cancers of the tongue and oropharynx particularly in males (Supplement S7). Comparison of HNC combined compared to the cancer registry data across the UK[43,44,45] shows lower but comparable trends (Supplement S8).

Overall IR was highest with increasing age peaking at 70-79 years old: 33.2 (31.9 to 34.4) per 100,000 person-years in CPRD GOLD with similar values in Aurum (28.5 [27.7 to 29.3]). For cancers of the oropharynx and tongue, IR was highest in those aged 60 to 69 years whereas cancers of the

medRxiv preprint doi: https://doi.org/10.1101/2024.12.05.24318538; this version posted December 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

larynx, hypopharynx and nasopharynx, overall IR was highest in those 70-79 years old. For cancers of the oral cavity, those aged 80-89 years had the highest overall IR whereas for the salivary gland, IR was highest in those 90 years and older. For cancers of the nasal cavity and sinus, IR were highest in those aged 70-79 years old for GOLD and highest in 80–89-year-olds in Aurum (Supplement S9).

For different age groups, those aged 30-39 and 90+, IR were relatively stable over time whereas IR for those aged 40-59 and 80-89 increased to 2006 before stabilising. IR for those aged 60-79 gradually increased over time. In GOLD, IR decreased in 2020 for those aged 40-69 and those 90 years and older (Figure 3). Stratifying by age and sex showed similar trends with higher and larger increases in IR over time for males (Supplement 10).

Annualised IR stratified by age group for each HNC subsite showed some similar trends to Figure 3 (Supplement 11). Cancers of the hypopharynx and oral cavity, IR remained relatively stable for 50–59-year-olds. For laryngeal cancer, those aged 50-69 years IR decreased over time whereas those aged 40-49 and 70-89 showed stable IR over time. For cancers of the oropharynx, salivary gland, and tongue, IR increased over time for all age groups (Supplement S12.1-S12.6).

Crude period prevalence stratified by calendar year, age, and sex.

PP for HNC in 2021 for GOLD was 0.13% (0.12% to 0.13%). PP in 2021 was 2.3-fold higher for males 0.18% [0.17% to 0.18%]) compared to females (0.08% [0.07% to 0.08%]). PP increased 3.6-fold from 2000 to 2021 in GOLD with fold increases of 3.3 and 3.7 for females and males respectively. PP in 2021 for HNC subsites was highest for cancer of the tongue, larynx, and oropharynx whereas PP was lowest for cancers of the nasal cavity and sinuses. Males had highest PP in 2021 (0.04%) for cancer of the larynx, tongue and oropharynx, whereas for females, cancer of the tongue, oral cavity and oropharynx had the highest PP in 2021 (0.02%) with similar trends in Aurum.

Annualised crude PP increased over time for HNC and subsites apart from hypopharynx, nasopharynx, and oral cavity where PP decreased from 2019 (Figure 3). Similar trends were seen for both databases with some exceptions. For Aurum, PP decreased from 2012 for hypopharyngeal cancer whereas for oral cavity, PP were stable from 2010 onwards with PP being stable from 2010 declining from 2015 for laryngeal cancer. Oropharyngeal cancer had the largest fold increase in PP across the study period for both databases. Sex stratification showed similar trends (supplement S13). Age stratification showed PP increased over time for those over 60 years of age for HNC. For

medRxiv preprint doi: https://doi.org/10.1101/2024.12.05.24318538; this version posted December 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

those aged 40-59, PP increased before stabilising from 2015. Stratification on sex and age showed males with higher prevalence from 40 years of age (supplement S14). Annualised PP stratified by age group for each HNC subsite showed some similar trends (Supplement S15), with exceptions for hypopharyngeal and laryngeal cancer (Supplement S16.1-16.7).

Survival with age, sex, and calendar year stratification

In GOLD there were 12,381 patients with 5,662 deaths (45.7% of patients) over the study period. Median survival for GOLD was 7.0 (6.6 - 7.5) years. Survival after one, five, and ten years after diagnosis was 81%, 56%, and 41% in GOLD with similar results in Aurum (Supplement S17).

Females had a longer median survival (8.2 years) compared to males (6.5 years) with better five-year survival (59.4%) compared to males (55.1%). Cancer of the hypopharynx had the lowest median survival of 3.4 years and 2.6 years in GOLD and Aurum (supplement S18). Cancer of the oral cavity had the longest median survival of 8.7 years however in Aurum, cancer of the salivary glands had the longest median survival (9.6 years).

Across the different subsites median survival was similar for both databases apart from cancer of the nasopharynx (6.1 years GOLD; 9.1 Aurum) and for cancers of salivary glands where survival between databases was greater than one year. Stratifying by sex, females showed slightly longer median survival for most subsites.

Overall, short, and long-term survival was similar for most subsites with one-, five- and ten-year survival of around 77-83%, 50-60% and 40-50% respectively apart from cancers of the hypopharynx and the nasal cavity and sinuses which was lower (Table 2). There were no sex differences in short- or long-term survival apart from cancer of the salivary gland where females had better short- and long-term term survival.

Table 2: Survival rates (and 95% confidence intervals) of head and neck cancer subsites from in CPRD GOLD (2000-2021) and CPRD Aurum (2000-2019) stratified by database.

	One-year survival (%)		Five-year survival (%)		Ten-year survival (%)	
Cancer	**CPRD GOLD**	**CPRD Aurum**	**CPRD GOLD**	**CPRD Aurum**	**CPRD GOLD**	**CPRD Aurum**
Hypopharynx	71.9 (68.7 - 75.1)	68.6 (65.8 - 71.5)	42.5 (38.8 - 46.5)	36.8 (33.7 - 40.2)	23.8 (20.0 - 28.4)	24.2 (21.1 - 27.7)
Larynx	82.7 (81.2 - 84.1)	82.5 (81.5 - 83.6)	55.9 (53.9 - 58.0)	56.5 (55.1 - 58.0)	38.8 (36.4 - 41.3)	40.5 (38.8 - 42.1)
Nasal Cavity & Sinus	73.4 (68.1 - 79.2)	75.9 (72.1 - 79.9)	50.2 (43.9 - 57.3)	47.6 (42.8 - 52.8)	34.5 (27.2 - 43.6)	34.3 (29.4 - 40.1)
Nasopharynx	77.4 (73.2 - 82.0)	80.9 (78.1 - 83.7)	52.1 (46.6 - 58.2)	59.0 (55.2 - 63.1)	37.8 (31.4 - 45.6)	47.5 (43.0 - 52.5)
Oral Cavity	84.7 (83.1 - 86.4)	82.8 (81.3 - 84.3)	62.0 (59.6 - 64.5)	59.8 (57.7 - 61.9)	46.0 (43.1 - 49.1)	43.5 (41.1 - 46.1)
Oropharynx	82.2 (80.6 - 83.8)	82.6 (81.5 - 83.8)	58.7 (56.4 - 61.0)	60.0 (58.3 - 61.7)	45.1 (42.3 - 48.1)	47.9 (45.8 - 50.0)
Salivary Gland	83.8 (81.2 - 86.6)	83.4 (81.5 - 85.3)	58.3 (54.4 - 62.5)	59.6 (56.9 - 62.4)	44.6 (39.7 - 50.0)	49.4 (46.3 - 52.8)
Tongue	78.6 (77.0 - 80.3)	80.7 (79.5 - 81.9)	54.6 (52.5 - 56.9)	57.3 (55.7 - 59.0)	42.3 (39.7 - 45.1)	44.6 (42.7 - 46.7)
All HNC subsites	81.0 (80.3 - 81.7)	81.3 (80.8 - 81.8)	56.4 (55.4 - 57.4)	56.8 (56.1 - 57.6)	41.3 (40.1 - 42.5)	42.7 (41.8 - 43.6)

medRxiv preprint doi: https://doi.org/10.1101/2024.12.05.24318538; this version posted December 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

medRxiv preprint doi: https://doi.org/10.1101/2024.12.05.24318538; this version posted December 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Stratification by calendar year in GOLD (supplement S19) showed short-term survival has not improved. However, oral cancer survival decreased from 89.8% (86.5-93.3) in 2000-2004 to 82.6% (78.9-86.4) in 2015-2019. Long term survival for HNC improved from 53.8% (51.4-56.3) in 2000-2004 to 58.7% (56.5-60.9) in 2015-2019. Nasopharyngeal cancer showed improvements from 32.8% (22.0-49.0) to 63.1% (52.4-76.0) in five-year survival with wide confidence intervals.

DISCUSSION

This study provides a comprehensive analysis of secular trends in HNC incidence, prevalence, and survival in the UK. Our analysis reveals notable increases in incidence and prevalence of HNC likely due to an ageing population, particularly among males and individuals aged 70–79 years old. Rises are primarily attributed to cancers of the tongue, salivary gland, and oropharynx. Laryngeal cancer appears to be stabilizing or, in some cases, decreasing. Patients diagnosed with hypopharyngeal, or nasal cavity and sinuses cancers exhibited poorest survival, compared to other HNC subsites. There has been a slight improvement in long-term survival rates across all HNC.

Our findings align with global trends, which have reported increases in HNC incidence and prevalence over recent decades[7, 12, 22], particularly for cancers of the tongue, oropharynx, and salivary glands [7]. While our IR for HNC broadly align with National Cancer Statistics across the UK [23,43,44,45], variations across countries, particularly for specific subsites, are likely due to different classification of subsites[56] as well as differing risk factors among populations [8, 24]. However, evidence suggests demographic profiles have remained unchanged over time despite rising oropharyngeal cancer incidence and stable or declining laryngeal cancer rates[57].

Reasons for increases in HNC cases are multifaceted and vary per subsite due to different etiologies and specific risk factors[8]. Smoking and alcohol consumption are the well-known behavioral risk factors associated with HNC, particularly among males, especially for oral cavity and larynx cancers which could have contributed to the rising numbers of HNC[13,25,26,27]. However, other studies have shown IR for these subsites falling in countries with a higher sociodemographic index, potentially related to decreasing smoking and alcohol consumption [28]. Therefore, increases in HNC cannot be attributed solely to smoking and alcohol consumption with an overall decline in these behaviors younger people in the UK[30,31]. Birth cohort effects likely play a role in the development of certain HNC

medRxiv preprint doi: https://doi.org/10.1101/2024.12.05.24318538; this version posted December 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

cancers due to varying exposure to risk factors over time on top of an increasing ageing population [7, 39].

Increasing HNC could be partly due to the increases in HPV infections [31] which has led to a notable increase in HPV-associated HNC, with rates for most non-HPV-related HNC remaining stable or declining[32]. However, despite a potential link between HPV and HNC subsites[54], one study found no change in the HPV-attributable HNC cases between 2002 and 2011, despite oropharyngeal cancer doubling, suggesting HPV plays a role but may not be entirely explained by changes in HPV infections [35.]

Survival estimates for HNC align with findings from the UK[8]and Europe[46] with cancers of the hypopharynx and nasal cavity and sinuses having lowest survival rates[47]. Poorer survival rates for these cancers can be attributed to late presentation, early lymph nodal metastasis, and high cancer recurrence[48] with most patients asymptomatic until diagnosis relative to other subsites[49].

Our study revealed no consistent improvements in short-term survival and only a slight improvement in long-term survival in the UK. This aligns with research indicating that survival rates for oral and pharyngeal cancer in the UK have not improved in recent decades[50]. As our study period covers up to 2021, small improvements in long term survival could be attributed to a larger reduction in smoking and alcohol before 2000's we well as better management using multi-disciplinary teams and advances in palliative care[51].

We used two large primary care databases with over 20 years of follow-up covering across the whole of the UK. CPRD GOLD covers practices from England, Wales, Scotland and Northern Ireland, while Aurum covers practices in England[15]. The use of both databases meant it was possible to compare results and increase generalizability of research findings. However, for some HNC subsites, there were some differences between databases likely due to geographical location[52]. Another strength of our study is the inclusion of a complete study population .in contrast, with cancer registries, potentially introducing bias[16].

Our study had several limitations. We used primary care data without linkage to cancer registries, which may have introduced misclassification or delays in recording. Our findings align with national cancer incidence trends, lending support to the external validity of our results. This consistency suggests that our data reflect broader population trends, despite the limitations of primary care-only

medRxiv preprint doi: https://doi.org/10.1101/2024.12.05.24318538; this version posted December 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

records[54]. Our use of primary care records precluded us from studying tumour histology, genetic mutations, staging or cancer therapies which can all impact survival. The anatomical proximity between the different sites of HNC often leads to erroneous classification of this type of cancer, which may result in misclassification. Finally, the HNC classification using SNOMED CT used in our study does not always align with the anatomical classification used by otolaryngologists [53]. For example, the tongue is divided anatomically into base of the tongue, and the anterior two-thirds which are classified as oropharynx, and oral cavity. However, unspecific SNOMED CT codes such as “Malignant neoplasm of tongue” meant we were unable to classify patients into either oral cavity or oropharyngeal HNC subsites meaning subsites may be underreported.

CONCLUSION

This study highlights increasing incidence and prevalence of HNC in the UK are largely driven by an aging population, while specific subsites are influenced by additional factors such as HPV infections. The lack of significant improvement in short-term survival and only modest gains in five-year survival emphasize the critical need for earlier diagnosis and tailored strategies to improve outcomes. Continued surveillance and research are essential to address the unique epidemiology of HNC subsites and guide prevention, diagnosis, and treatment efforts.

CONTRIBUTIONS

Conceptualization; DN, DPA, EB

Data harmonisation and data quality assessment; AD, WYM

Formal analysis; DN

Funding acquisition; DPA

Supervision; DPA, EB, FXA-J

Interpretation of results: All authors

Roles/Writing - original draft: AMD, DN

Writing - review & editing: All authors

FUNDING

This activity under the European Health Data & Evidence Network (EHDEN) and OPTIMA has received funding from the Innovative Medicines Initiative 2 (IMI2) Joint Undertaking under grant agreement No 806968 and No. 101034347 respectively. IMI2 receives support from the European

medRxiv preprint doi: https://doi.org/10.1101/2024.12.05.24318538; this version posted December 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Union's Horizon 2020 research and innovation programme and European Federation of Pharmaceutical Industries and Associations (EFPIA). The sponsors of the study did not have any involvement in the writing of the manuscript or the decision to submit it for publication. Additionally, there was partial support from the Oxford NIHR Biomedical Research Centre. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

CONFLICTS OF INTEREST

Professor Daniel Prieto-Alhambra research group has received research grants from the European Medicines Agency, from the Innovative Medicines Initiative, from Amgen, Chiesi, and from UCB Biopharma; and consultancy or speaker fees from Astellas, Amgen, Astra Zeneca, and UCB Biopharma. All other authors declare no conflicts of interest.

DATA AVAILIABILITY STATEMENT

This study is based in part on data from the Clinical Practice Research Datalink (CPRD) obtained under licence from the UK Medicines and Healthcare products Regulatory Agency. The data is provided by patients and collected by the NHS as part of their care and support. The interpretation and conclusions contained in this study are those of the author/s alone. Patient level data used in this study was obtained through an approved application-the CPRD (application number 22_001843) and is only available following an approval process-safeguard the confidentiality of patient data. Details on how to apply for data access can be found at https://cprd.com/data-access.

REFERENCES

1. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021;71(3):209-249.
2. Thomas SJ, Penfold CM, Waylen A, Ness AR. The changing aetiology of head and neck squamous cell cancer: a tale of three cancers? Clin Otolaryngol. 2018;43(5):999-1003.
3. Hashibe M, Brennan P, Chuang SC, et al. Interaction between tobacco and alcohol use and the risk of head and neck cancer: pooled analysis in the International Head and Neck Cancer Epidemiology Consortium. Cancer Epidemiol Biomarkers Prev. 2009;18(2):541-550.
4. Bravi F, Lee YA, Hashibe M, et al; INHANCE Consortium investigators. Lessons learned from the INHANCE consortium: an overview of recent results on head and neck cancer. Oral Dis. 2021;27(1):73-93. doi:10.1111/odi.13502

medRxiv preprint doi: https://doi.org/10.1101/2024.12.05.24318538; this version posted December 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

5. Roman BR, Aragones A. Epidemiology and incidence of HPV-related cancers of the head and neck. J Surg Oncol. 2021;124(6):920-922. doi:10.1002/jso.26687
6. Bravi F, Lee Y-C A, Hashibe M et al. Lessons learned from the INHANCE consortium: An overview of recent results on head and neck cancer. Oral Dis 2021; 27: 73-93.
7. McCarthy CE, Field JK, Rajlawat BP, Field AE, Marcus MW. Trends and regional variation in the incidence of head and neck cancers in England: 2002 to 2011. Int J Oncol. 2015;47:204-210.
8. Gormley M, Creaney G, Schache A, Ingarfield K, Conway DI. Reviewing the epidemiology of head and neck cancer: definitions, trends, and risk factors. Br Dent J. 2022;233(9):780-786. doi:10.1038/s41415-022-5166-x
9. Global Burden of Disease Cancer Collaboration. Global, Regional, and National Cancer Incidence, Mortality, Years of Life Lost, Years Lived With Disability, and Disability-Adjusted Life-years for 32 Cancer Groups, 1990 to 2015: A Systematic Analysis for the Global Burden of Disease Study. JAMA Oncol 2017; 3: 524-548.
10. Mehanna H, Beech T, Nicholson T, et al. Prevalence of human papillomavirus in oropharyngeal and nonoropharyngeal head and neck cancer: systematic review and meta-analysis of trends by time and region. Head Neck. 2013;35(6):747-755.
11. Pulte D, Brenner H. Changes in survival in head and neck cancers in the late 20th and early 21st century: a period analysis. Oncologist. 2010;15(9):994-1001. doi:10.1634/theoncologist.2009-0289
12. Creaney G, McMahon AD, Ross AJ, Bhatti LA, Paterson C, Conway DI. Head and neck cancer in the UK: what was the stage before COVID-19? UK cancer registries analysis (2011-2018). Br Dent J. 2022;233(9):787-793. doi:10.1038/s41415-022-5151-4
13. Bosetti C, Carioli G, Santucci C, et al. Global trends in oral and pharyngeal cancer incidence and mortality. *Int J Cancer*. 2020;147(4):1040-1049. doi:10.1002/ijc.32871
14. Warnakulasuriya S. Global epidemiology of oral and oropharyngeal cancer. *Oral Oncol*. 2009;45(4-5):309-316. doi:10.1016/j.oraloncology.2008.06.002
15. Herrett E, Gallagher AM, Bhaskaran K, et al. Data resource profile: clinical practice research datalink (CPRD). *Int J Epidemiol*. 2015;44(3):827-836.
16. Voss EA, Makadia R, Matcho A, et al. Feasibility and utility of applications of the common data model to multiple, disparate observational health databases. *J Am Med Inform Assoc*. 2015;22(3):553-564.
17. Hripcsak G, Duke JD, Shah NH, et al. Observational health data sciences and informatics (OHDSI): opportunities for observational researchers. *Stud Health Technol Inform*. 2015;216:574-578.
18. Schuemie M, Ryan P, Weaver J. Cohort diagnostics: diagnostics for OHDSI studies. 2020.
19. Eurostat Taskforce. Revision of the European Standard Population—Report of Eurostat's task force. 2013. https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/ks-ra-13-028. Accessed August 2, 2023.

medRxiv preprint doi: https://doi.org/10.1101/2024.12.05.24318538; this version posted December 6, 2024. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

20. Raventós B, et al. IncidencePrevalence: An R package to calculate population-level incidence rates and prevalence using the OMOP common data model. *Pharmacoepidemiol Drug Saf*. 2023;1-11.
21. Aupérin A. Epidemiology of head and neck cancers: an update. *Curr Opin Oncol*. 2020;32(3):178-186. doi:10.1097/CCO.0000000000000629
22. Kawakita D, Oze I, Iwasaki S, et al. Trends in the incidence of head and neck cancer by subsite between 1993 and 2015 in Japan. *Cancer Med*. 2022;11(6):1553-1560. doi:10.1002/cam4.4539
23. Cancer Research UK. Head and neck cancers incidence statistics. https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/head-and-neck-cancers/incidence. Accessed August 13, 2024.
24. Petersen JF, Timmermans AJ, van Dijk BAC, et al. Trends in treatment, incidence, and survival of hypopharynx cancer: a 20-year population-based study in the Netherlands. *Eur Arch Otorhinolaryngol*. 2018;275(1):181-189. doi:10.1007/s00405-017-4766-6
25. Johnson DE, Burtness B, Leemans CR, et al. Head and neck squamous cell carcinoma. *Nat Rev Dis Primers*. 2020;6(1):92. doi:10.1038/s41572-020-00224-3
26. Bagnardi V, Rota M, Botteri E, et al. Alcohol consumption and site-specific cancer risk: a comprehensive dose-response meta-analysis. *Br J Cancer*. 2015;112(4):580-593. doi:10.1038/bjc.2014.579
27. Koo HY, Han K, Shin DW, et al. Alcohol drinking pattern and risk of head and neck cancer: A nationwide cohort study. *Int J Environ Res Public Health*. 2021;18(21):11204. doi:10.3390/ijerph182111204
28. Mehanna H, Paleri V, West CM, Nutting C. Head and neck cancer—Part 1: Epidemiology, presentation, and prevention. *BMJ*. 2010;341
29. GBD 2016 Alcohol Collaborators. Alcohol use and burden for 195 countries and territories, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet*. 2018;392(10152):1015-1035. doi:10.1016/S0140-6736(18)31310-2
30. Hardie I, Sasso A, Holmes J, Meier PS. Understanding changes in the locations of drinking occasions in Great Britain: An age-period-cohort analysis of repeat cross-sectional market research data, 2001-2019. *Drug Alcohol Rev*. 2023;42(1):105-118. doi:10.1111/dar.13562
31. GBD 2019 Tobacco Collaborators. Spatial, temporal, and demographic patterns in prevalence of smoking tobacco use and attributable disease burden in 204 countries and territories, 1990-2019: a systematic analysis from the Global Burden of Disease Study 2019. *Lancet*. 2021;397(10292):2337-2360. doi:10.1016/S0140-6736(21)01169-7
32. Bruni L, Albero G, Rowley J, et al. Global and regional estimates of genital human papillomavirus prevalence among men: a systematic review and meta-analysis. Lancet Glob Health. 2023;11(9)
33. Menezes FDS, Fernandes GA, Antunes JLF, et al. Global incidence trends in head and neck cancer for HPV-related and -unrelated subsites: A systematic review of population-based studies. Oral Oncol. 2021;115:105177. doi:10.1016/j.oraloncology.2020.105177

medRxiv preprint doi: https://doi.org/10.1101/2024.12.05.24318538; this version posted December 6, 2024. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

34. Deshmukh AA, Damgacioglu H, Georges D, et al. Global burden of HPV-attributable squamous cell carcinoma of the anus in 2020, according to sex and HIV status: A worldwide analysis. Int J Cancer. 2023;152(3):417-428. doi:10.1002/ijc.34269
35. Schache AG, Powell NG, Cuschieri KS, et al. HPV-related oropharyngeal cancer in the United Kingdom: an evolution in understanding of disease etiology. *Cancer Res*. 2016;76:6598-6606.
36. Aldalwg MAH, Brestovac B. Human Papillomavirus Associated Cancers of the Head and Neck: An Australian Perspective. *Head Neck Pathol*. 2017;11(3):377-384. doi:10.1007/s12105-017-0780-7
37. Chow LQM. Head and Neck Cancer. *N Engl J Med*. 2020;382(1):60-72. doi:10.1056/NEJMra1715715
38. Marziliano A, Teckie S, Diefenbach MA. Alcohol-related head and neck cancer: Summary of the literature. *Head Neck*. 2020;42(4):732-738. doi:10.1002/hed.26023
39. Hertrampf K, Wiltfang J, Katalinic A, Timm O, Wenz HJ. Trends in incidence, tumour sites, and tumour stages of oral and pharyngeal cancer in Northern Germany. *J Cancer Res Clin Oncol*. 2012;138:431-437.
40. Langevin SM, Grandis JR, Taioli E. Female hormonal and reproductive factors and head and neck squamous cell carcinoma risk. *Cancer Lett*. 2011;310(2):216-221. doi:10.1016/j.canlet.2011.07.007
41. Therneau T. A package for survival analysis in R. R package version 3.7-0. *CRAN*. 2024. https://CRAN.R-project.org/package=survival
42. National Cancer Registration and Analysis Service. Cancer Incidence and Mortality. https://www.cancerdata.nhs.uk/incidence_and_mortality. Accessed August 13, 2024.
43. Cancer incidence in Scotland—cancer incidence and prevalence in Scotland to December 2019. Public Health Scotland. https://publichealthscotland.scot/publications/cancer-incidence-in-scotland/cancer-incidence-in-scotland-cancer-incidence-and-prevalence-in-scotland-to-december-2019. Accessed August 13, 2024.
44. Cancer Incidence in Wales, 2002-2019. Public Health Wales. Accessed August 13, 2024. 2019/cancersurvivalinwalestechnicalguide2002-2019pdf/https://phw.nhs.wales/services-and-teams/
45. 1993-2020. Northern Ireland Cancer Registry. Accessed August 13, 2024. https://www.qub.ac.uk/research-centres/nicr/CancerInformation/official-statistics/BySite/
46. Guizard AV, Dejardin OJ, Launay LC, et al. Diagnosis and management of head and neck cancers in a high-incidence area in France: A population-based study. Medicine. 2017;96(26) doi:10.1097/MD.0000000000007285
47. Du E, Mazul AL, Farquhar D, et al. Long-term survival in head and neck cancer: impact of site, stage, smoking, and human papillomavirus status. Laryngoscope. 2019;129(11):2506-2513. doi:10.1002/lary.27807
48. Hall SF, Groome PA, Irish J, O'Sullivan B. The natural history of patients with squamous cell carcinoma of the hypopharynx. Laryngoscope. 2008;118(8):1362-1371.

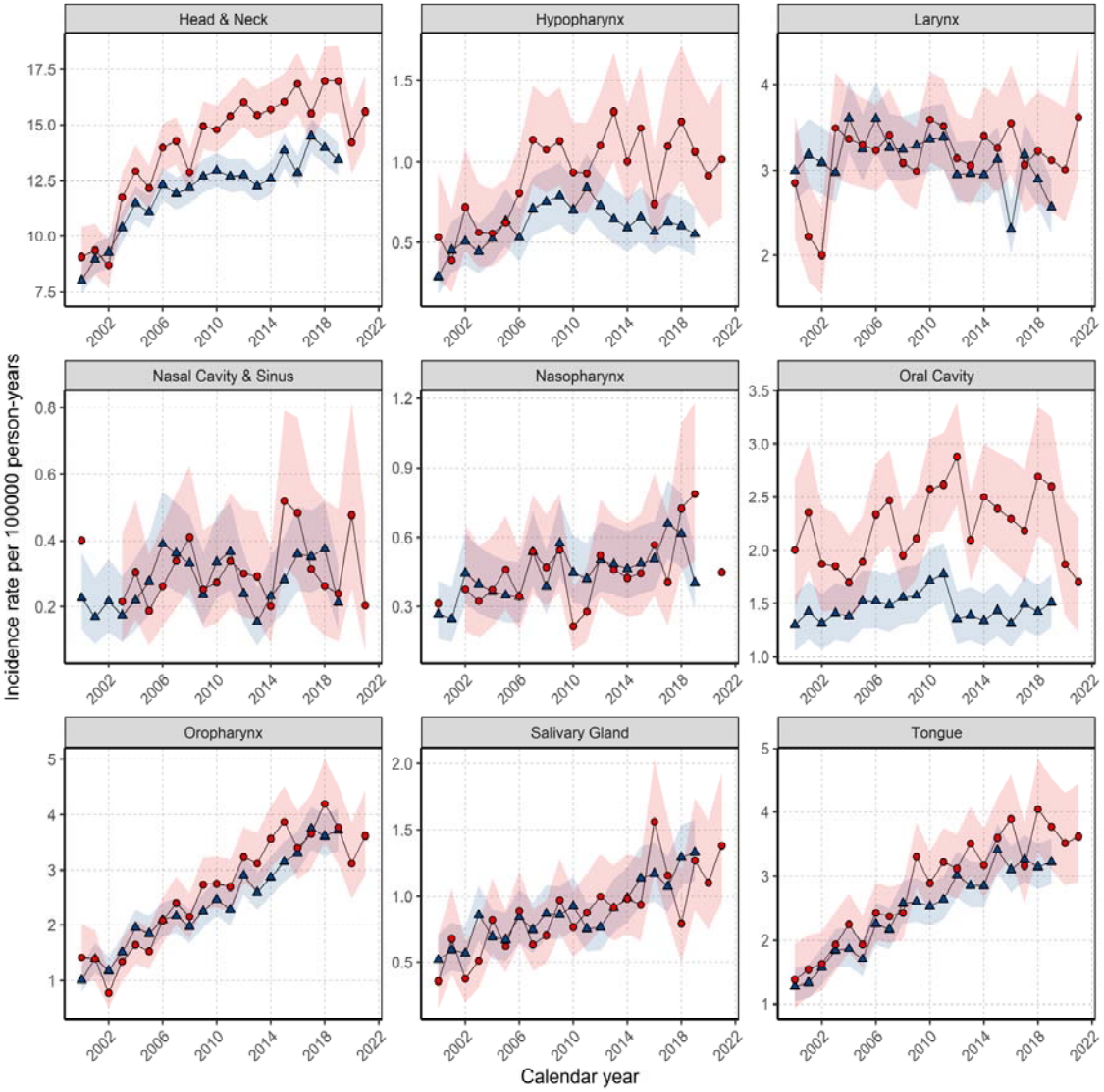
medRxiv preprint doi: https://doi.org/10.1101/2024.12.05.24318538; this version posted December 6, 2024. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

49. Bugra Cengiz A, Uyar M, Comert E, et al. Sinonasal tract malignancies: prognostic factors and surgery outcomes. Iran Red Crescent Med J. 2013;15(12) doi:10.5812/ircmj.14118
50. Warnakulasuriya S. Global epidemiology of oral and oropharyngeal cancer. Oral Oncol. 2009;45(4-5):309-316. doi:10.1016/j.oraloncology.2008.06.002
51. Abdulrahman GO Jr. The effect of multidisciplinary team care on cancer management. Pan Afr Med J. 2011;9:20. doi:10.4314/pamj.v9i1.71195
52. Mayor. S. Survival rates with head and neck cancers vary widely, audit finds. BMJ. 2015;351 doi:10.1136/bmj.h4758.
53. Burgun A, Bodenreider O, Mougin F. Classifying diseases with respect to anatomy: a study in SNOMED CT. AMIA Annu Symp Proc. 2005;2005:91-95. PMID: 16779008; PMCID: PMC1560776.
54. Strongman H, Williams R, Bhaskaran K. What are the implications of using individual and combined sources of routinely collected data to identify and characterise incident site-specific cancers? A concordance and validation study using linked English electronic health records data. BMJ Open. 2020;10
55. Kim, S.I., Lee, J.W., Eun, YG. et al. A SEER-based analysis of trends in HPV-associated oropharyngeal squamous cell carcinoma. Infect Agents Cancer 19, 29 (2024). https://doi.org/10.1186/s13027-024-00592
56. *The Larynx and Its Subsites in Head and Neck Cancer Staging.* American Academy of Otolaryngology-Head and Neck Surgery. Accessed November 6, 2024. https://www.entnet.org:contentReference[oaicite:1]{index=1}
57. Smith, C.D.L., McMahon, A.D., Purkayastha, M. et al. Head and neck cancer incidence is rising but the sociodemographic profile is unchanging: a population epidemiological study (2001–2020). BJC Rep 2, 71 (2024). https://doi.org/10.1038/s44276-024-00089-z
58. Oxford Pharma Epidemiology. EHDEN Cancer Incidence and Prevalence Study. GitHub repository. Published . Accessed November 21, 2024. https://github.com/oxford-pharmacoepi/EHDENCancerIncidencePrevalence

medRxiv preprint doi: https://doi.org/10.1101/2024.12.05.24318538; this version posted December 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

FIGURES

Figure 1: Crude annualised incidence rates for HNC and subsites stratified by database (2004 was the initiation of the Quality Outcome Framework in the UK).



medRxiv preprint doi: https://doi.org/10.1101/2024.12.05.24318538; this version posted December 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

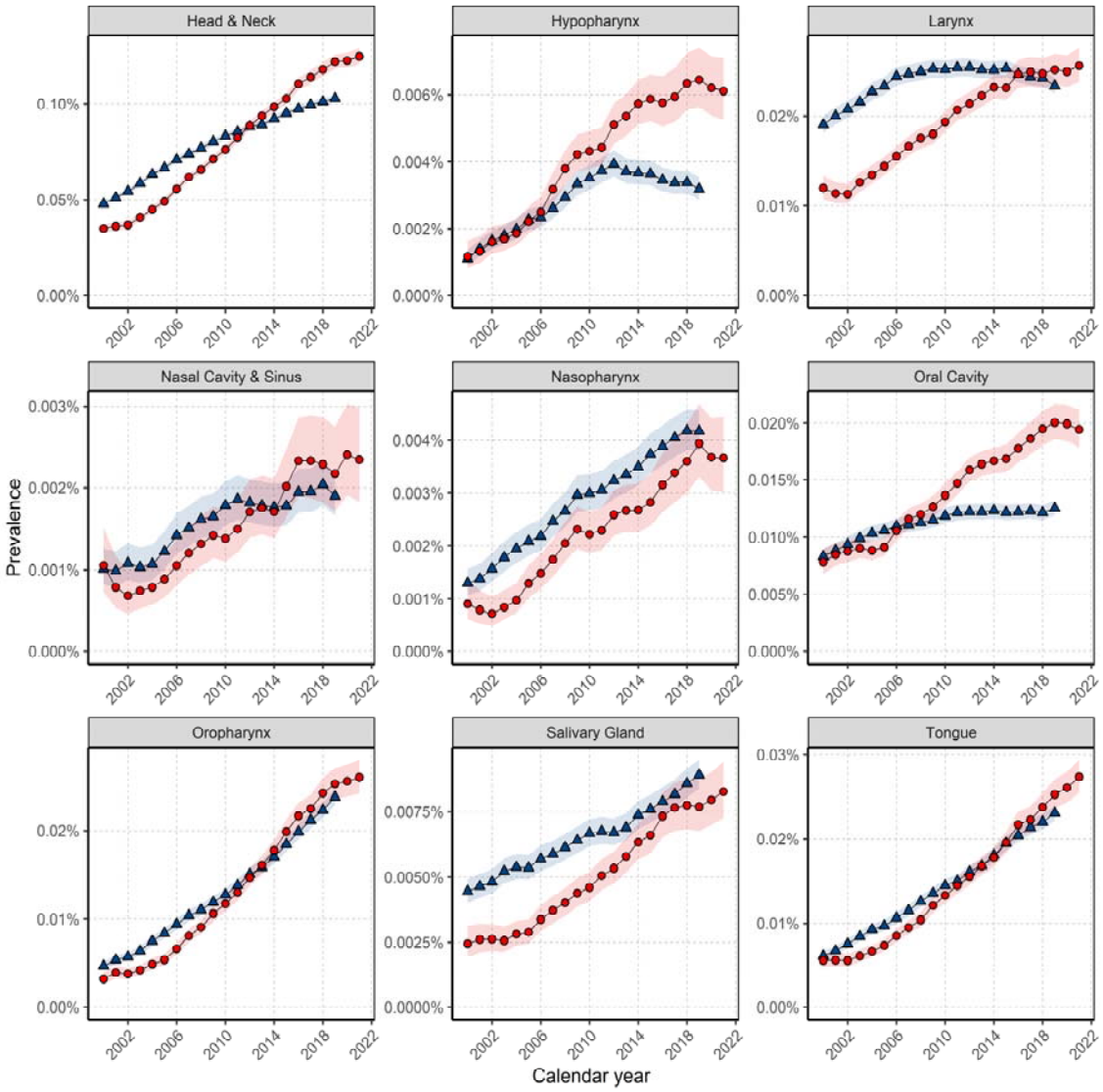



Figure 3: Annualised crude period prevalence for HNC and subsites.