Check for updates

DATA NOTE

The HEADSpAcE Data Center: An international data and biospecimen resource for head and neck cancer translational research

[version 1; peer review: awaiting peer review]

Shama Virani[1], Laia Alemany[2-4], Wolfgang Ahrens[5], Shaymaa AlWaheidi[6], Lidia Arantes[7], Aude Bardot[8], Paul Brennan[1], Cristina Canova[9], Amelie Chabrier[1], David I Conway[10], Grant Creaney[10], Mauricio Cuello[11], Cecilia Cuffini[12], Maria Paula Curado[13], Ana Carolina de Carvalho[1], Evandro de Souza[14], Brenda Diergaarde[15,16], Tom Dudding[17], Valerie Gaborieau[1], David Hayes[18], Claire Healy[19], Ivana Holcatova[20,21], Rayjean Hung[22,23], Jens Klussmann[24,25], Pagona Lagiou[26], Areti Lagiou[27], Geoffrey Liu[28], Clovis Lopes Pinto[13], Silvia Lopez de Blanc[29], Louise McNeilage[1], Laura Mendoza[30], Jose Carlos Oliveira[31], Andrew Olshan[32], Hernan Ortiz[33], Sandra Perdomo[1], Shahid Pervez[34], Jerry Polesel[35], Miranda Pring[36], Luis Felipe Ribeiro Pinto[37], Lorenzo Richardi[38,39], Max Robinson[40], Paula Rodriguez-Urrego[41], Pierre-Eric Roux[42], Pierre Saintigny[43], Stacey Santi[44], Deborah Saunders[45], Marta Vilensky[46], Sandra von Zeidler[47], Tim Waterboer[48], Kazem Zendehdel[49], Ariana Znaor[8], HEADSpAcE Consortium

[1]Genomic Epidemiology Branch, International Agency for Research on Cancer, Lyon, 69007, France
[2]Unit of Molecular Epidemiology and Genetics (UNIC EMG), Cancer Epidemiology Research Programme, Institut Catala d'Oncologia, L'Hospitalet de Llobregat, Catalonia, Spain
[3]Biomedical Research Networking Center for Epidemiology and Public Health, Instituto de Salud Carlos III, Madrid, Spain
[4]Bellvitge Biomedical Research Institute, Hospital General de l'Hospitalet, L'Hospitalet de Llobregat, Catalonia, Spain
[5]Leibniz Institute for Prevention Research and Epidemiology-Bremen, Bremen, Germany
[6]Early Detection, Prevention Infections Branch, International Agency for Research on Cancer, Lyon, France
[7]Molecular Oncology Research Center, Hospital de Cancer de Barretos, Barretos, State of São Paulo, Brazil
[8]Cancer Surveillance Branch, International Agency for Research on Cancer, Lyon, France
[9]Department of Cardiac Thoracic Vascular Sciences and Public Health, University of Padua Department of Cardiac Thoracic and Vascular Science and Public Health, Padua, Veneto, Italy
[10]University of Glasgow School of Medicine Dentistry and Nursing, Glasgow, Scotland, UK
[11]Servicio de Oncología Clínica Hospital de Clínicas, Universidad de la Republica Facultad de Medicina, Montevideo, Montevideo Department, Uruguay
[12]Instituto de Virología "Dr. J. M. Vanella", Universidad Nacional de Cordoba Facultad de Ciencias Medicas, Córdoba, Cordoba, Argentina
[13]Group of Epidemiology and Statistics on Cancer, Fundacao Antonio Prudente, São Paulo, State of São Paulo, Brazil
[14]Head and Neck Surgery Division, Hospital Santa Rita de Cássia, Associação Feminina de Educação e Combate ao Câncer, Vitoria, Brazil
[15]UPMC Hillman Cancer Center, Pittsburgh, Pennsylvania, USA
[16]Department of Human Genetics, University of Pittsburgh School of Public Health, Pittsbirgh, USA
[17]University of Bristol School of Oral and Dental Sciences, Bristol, England, UK

[18]The University of Tennessee Health Science Center Department of Medicine, Memphis, Tennessee, USA
[19]School of Dental Science, Trinity College Dublin Faculty of Health Sciences, Dublin, Leinster, Ireland
[20]Institute of Hygiene and Epidemiology, Charles University First Faculty of Medicine, Prague, Prague, Czech Republic
[21]Department of Oncology, Motol University Hospital, Prague, Prague, Czech Republic
[22]Prosserman Centre for Population Health Research, Lunenfeld-Tanenbaum Research Institute, Sinai Health System, Toronto, Canada
[23]Division of Epidemiology, University of Toronto Dalla Lana School of Public Health, Toronto, Ontario, Canada
[24]Department of Otorhinolaryngology, Head and Neck Surgery, Medical Faculty of the University of Cologne, Cologne, North Rhine-Westphalia, Germany
[25]Center for Molecular Medicine Cologne, Universitat zu Koln Medizinische Fakultat, Cologne, North Rhine-Westphalia, Germany
[26]Department of Preventive Medicine, Epidemiology and Medical Statistics, National and Kapodistrian University of Athens School of Medicine, Athens, Greece
[27]Department of Public and Community Health, University of West Attica, School of Public Health, Athens, Greece
[28]Princess Margaret Cancer Centre, University Health Network, Toronto, Ontario, Canada
[29]Oral Pathology, School of Dentistry, National University of Cordoba, Cordoba, Argentina
[30]Instituto de Investigaciones en Ciencias de la Salud, Universidad Nacional de Asuncion, San Lorenzo, Central Department, Paraguay
[31]Hospital Câncer Araújo Jorge, Goiania, Goias, Brazil
[32]Department of Epidemiology, Gillings School of Global Public Health, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA
[33]Hospital de Clínicas, National University of Asunción School of Medical Sciences, Asunción, Capital District, Paraguay
[34]Department of Pathology and Laboratory Medicine, The Aga Khan University Hospital, Karachi, Sindh, Pakistan
[35]Unit of Cancer Epidemiology, Centro di Riferimento Oncologico IRCCS, Aviano, Friuli-Venezia Giulia, Italy
[36]University of Bristol Faculty of Health Sciences, Bristol, England, UK
[37]Instituto Nacional de Câncer, Rio de Janiero, Brazil
[38]University of Turin Department of Medical Sciences, Turin, Piedmont, Italy
[39]Cancer Epidemiology Unit, University Hospital Città della Salute e della Scienza, Turin, Italy
[40]The Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle, UK
[41]Pathology and Laboratory Department, University Hospital of the Fundacion Santa Fe de Bogota, Bogotá, Bogota, Colombia
[42]Department of Surgical Oncology, Centre Leon Berard, Lyon, Auvergne-Rhône-Alpes, France
[43]Department of Medical Oncology, Centre Leon Berard, Lyon, Auvergne-Rhône-Alpes, France
[44]Clinical Oncology, Health Sciences North Research Institute, Sudbury, Canada
[45]Department of Dental Oncology, Northern Ontario School of Medicine University, Sudbury, Canada
[46]Instituto de Oncologia Angel H Roffo, University of Buenos Aires, Buenos Aires, Autonomous City of Buenos Aires, Argentina
[47]Department of Pathology, Federal University of Espirito Santo, Vitoria, State of Espírito Santo, Brazil
[48]Division of Infections and Cancer Epidemiology, German Cancer Research Center, Heidelberg, Baden-Württemberg, Germany
[49]Tehran University of Medical Sciences Cancer Research Center, Tehran, Tehran Province, Iran

v1 **First published:** 07 Oct 2026, **6**:400
https://doi.org/10.12688/openreseurope.24719.1
Latest published: 07 Oct 2026, **6**:400
https://doi.org/10.12688/openreseurope.24719.1

Abstract

Head and neck cancer (HNC) comprises a heterogeneous group of malignancies with substantial variation in incidence, risk factors, and outcomes across populations. Existing studies are often limited by regional scope or lack integration of epidemiological, clinical, and molecular data, restricting the ability to investigate determinants of disease risk, diagnosis, and survival on a global scale.

The HEADSpAcE Data Center was established to address these gaps by

Open Peer Review

Approval Status *AWAITING PEER REVIEW*

Any reports and responses or comments on the article can be found at the end of the article.

integrating data and biospecimens from multiple international studies and consortia across North America, South America, Europe, and the Eastern Mediterranean under a shared governance framework that defines future access and use of the resource as has been agreed upon by participating principal investigators. The resource currently includes over 27,000 participants from 18 countries, encompassing detailed demographic, lifestyle, clinical, and follow-up/outcome data, with linked biospecimens and molecular profiling. Cancer cases included are from the oral cavity, oropharynx, larynx, and hypopharynx. Participant recruitment spans the periods from 1991 to 2025.

The Data Center incorporates standardized variables on risk factors (e.g., tobacco, alcohol, HPV), tumor characteristics, and disease outcomes, as well as detailed information on diagnostic pathways from the most recent recruitment. Molecular data include HPV-related assays, germline genotyping, targeted and/or genome-wide tumor sequencing, and transcriptomic tumor profiling. Biospecimens include blood-derived samples, oral samples and tumor tissues, with availability varying by region and study.

By combining multi-level data within a harmonized framework, the HEADSpAcE Data Center provides a platform for large-scale analyses of HNC risk, diagnosis, and survival across populations. The resource is open to new centers wishing to contribute data and biospecimens to expand representation across regions and populations. Data access and contribution procedures are available at www.headspace.iarc.who.int.

Keywords
HEADSpAcE, head and neck cancers, oral cancers, oropharynx cancers, HPV-related cancers, larynx cancers, data resource, international, standardized data collection

H2020 This article is included in the Horizon 2020 gateway.

Corresponding author: Shama Virani (viranis@iarc.who.int)

Author roles: Virani S: Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Methodology, Project Administration, Resources, Supervision, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Alemany L**: Data Curation, Funding Acquisition, Writing – Review & Editing; **Ahrens W**: Data Curation, Writing – Review & Editing; **AlWaheidi S**: Project Administration, Writing – Review & Editing; **Arantes L**: Data Curation, Writing – Review & Editing; **Bardot A**: Software, Writing – Review & Editing; **Brennan P**: Funding Acquisition, Writing – Review & Editing; **Canova C**: Data Curation, Writing – Review & Editing; **Chabrier A**: Software, Writing – Review & Editing; **Conway DI**: Data Curation, Funding Acquisition, Writing – Review & Editing; **Creaney G**: Data

Curation, Writing – Review & Editing; **Cuello M**: Data Curation, Writing – Review & Editing; **Cuffini C**: Data Curation, Writing – Review & Editing; **Curado MP**: Data Curation, Funding Acquisition, Writing – Review & Editing; **de Carvalho AC**: Data Curation, Funding Acquisition, Writing – Review & Editing; **de Souza E**: Data Curation, Writing – Review & Editing; **Diergaarde B**: Data Curation, Funding Acquisition, Writing – Review & Editing; **Dudding T**: Data Curation, Funding Acquisition, Writing – Review & Editing; **Gaborieau V**: Data Curation, Project Administration, Software, Writing – Review & Editing; **Hayes D**: Data Curation, Funding Acquisition, Writing – Review & Editing; **Healy C**: Data Curation, Writing – Review & Editing; **Holcatova I**: Data Curation, Writing – Review & Editing; **Hung R**: Data Curation, Funding Acquisition, Writing – Review & Editing; **Klussmann J**: Data Curation, Writing – Review & Editing; **Lagiou P**: Data Curation, Writing – Review & Editing; **Lagiou A**: Data Curation, Writing – Review & Editing; **Liu G**: Data Curation, Funding Acquisition, Writing – Review & Editing; **Lopes Pinto C**: Data Curation, Writing – Review & Editing; **Lopez de Blanc S**: Data Curation, Writing – Review & Editing; **McNeilage L**: Project Administration, Writing – Review & Editing; **Mendoza L**: Data Curation, Writing – Review & Editing; **Oliveira JC**: Data Curation, Writing – Review & Editing; **Olshan A**: Data Curation, Funding Acquisition, Writing – Review & Editing; **Ortiz H**: Data Curation, Writing – Review & Editing; **Perdomo S**: Data Curation, Funding Acquisition, Writing – Review & Editing; **Pervez S**: Data Curation, Writing – Review & Editing; **Polesel J**: Data Curation, Writing – Review & Editing; **Pring M**: Data Curation, Writing – Review & Editing; **Ribeiro Pinto LF**: Data Curation, Funding Acquisition, Writing – Review & Editing; **Richardi L**: Data Curation, Funding Acquisition, Writing – Review & Editing; **Robinson M**: Data Curation, Writing – Review & Editing; **Rodriguez-Urrego P**: Data Curation, Writing – Review & Editing; **Roux PE**: Data Curation, Writing – Review & Editing; **Saintigny P**: Data Curation, Writing – Review & Editing; **Santi S**: Data Curation, Writing – Review & Editing; **Saunders D**: Data Curation, Writing – Review & Editing; **Vilensky M**: Data Curation, Writing – Review & Editing; **von Zeidler S**: Data Curation, Writing – Review & Editing; **Waterboer T**: Data Curation, Funding Acquisition, Writing – Review & Editing; **Zendehdel K**: Data Curation, Writing – Review & Editing; **Znaor A**: Data Curation, Writing – Review & Editing;

Competing interests: No competing interests were disclosed.

Grant information: The HEADSpAcE consortium received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 825771 (HEADSpAcE project), Grant 2020/09088-1, São Paulo Research Foundation (FAPESP), Grant 2020/08805-1, São Paulo Research Foundation (FAPESP), Grant 404227/2019-8, Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Grant 33.654.831/0001-36, Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) The VOYAGER consortium was funded by US National Institute of Dental and Craniofacial Research (NIDCR) grant R01DE025712 (PIs: Brennan, Diergaarde, and Hayes: The role of germline and somatic DNA mutations in oral and oropharyngeal cancers). Genotyping using the OncoArray and the All of Us array was performed at the Center for Inherited Disease (CIDR) and funded by NIDCR 1X01HG007780-0 and NIDCR/NCI X01HG010743, respectively. The Alcohol-Related Cancers and Genetic Susceptibility Study in Europe (ARCAGE) was funded by the European Commission's fifth frame-work program (QLK1-2001-00182), the Italian Association for Cancer Research, Compagnia di San Paolo/FIRMS, Region Piemonte and Padova University (CPDA057222). The Carolina Head and Neck Cancer Epidemiology (CHANCE) study was supported in part by the National Cancer Institute (R01CA90731). The Head and Neck 5000 study was a component of independent research funded by the National Institute for Health Research (NIHR) under its Programme Grants for Applied Research scheme (RP-PG-0707-10034). The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health. Core funding was also provided through awards from Above and Beyond, University Hospitals Bristol and Weston Research Capability Funding and the NIHR Senior Investigator award to Professor Andy Ness. Human papilloma virus (HPV) serology was supported by a Cancer Research UK Programme Grant, the Integrative Cancer Epidemiology Programme (grant number: C18281/A19169). The University of Pittsburgh head and neck cancer case-control study was supported by US National Institutes of Health grants P50CA097190 and P30CA047904. The Toronto MSH-PMH study was supported by the Canadian Cancer Society Research Institute and the Princess Margaret Head & Neck Translational Research Program, with philanthropic funds from Joe's Team and the Wharton, Elia, Riley, and Tozer families. G.L. is funded by the M. Qasim Chair in Translational Research and the Lusi Wong Foundation Fund. The Head and Neck Intensive Clinical Tobocca Intervention (HN-iCTI) (Sudbury, ON) received research grant funding from the Health Sciences North Foundation (DPS) Barcelona study was supported by grants from the Instituto de Salud Carlos III-ISCIII (Spanish Government) cofounded by FEDER funds / European Regional Development Fund (ERDF) - a way to build Europe (References: PI1102096, PI1401918, PI1500500, PI1501205, RD12/0036/0056, CIBERESP, CIBERONC), Agència de Gestió d'Ajuts Universitaris i de Recerca (2014SGR756; 2014SGR1077), Beca de recerca clínica 2016 de l'Acadèmia de Ciències Mèdiques de Catalunya i Balears, Beca PERIS-2016-2020 Pla Estratègic de Recerca en Investigacions Sanitàries (SLT002/16/00404), Asociación Española Contra el Cáncer (personal grant to LA), Rio Hortega-SEOM (ISCIII-SpanishSociety of Medical Oncology) (personal grant to MT), Ayudas Merck Serono-Fundación Salud 2000 de investigación 2015 and from Sanofi Pasteur MSD and Merck & Co, Inc., who had no role in the data collection, analysis or interpretation of the results.The Centro di Riferimento Oncologico di Aviano (CRO), IRCCS was supported by Italian Ministry of Health – Ricerca Corrente.

Copyright: © 2026 Virani S *et al*. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Virani S, Alemany L, Ahrens W *et al.* **The HEADSpAcE Data Center: An international data and biospecimen resource for head and neck cancer translational research [version 1; peer review: awaiting peer review]** Open Research Europe 2026, **6**:400 https://doi.org/10.12688/openreseurope.24719.1

First published: 07 Oct 2026, **6**:400 https://doi.org/10.12688/openreseurope.24719.1

HEADSpAcE Data Center: Rationale and Scope

Head and neck cancer (HNC) comprises a heterogeneous group of malignancies arising in the oral cavity, oropharynx, larynx, and hypopharynx. Together, these cancers account for more than 845,493 new cases and 392,298 deaths annually worldwide, representing about 4% of the global cancer burden,[1] and exhibit substantial geographic variation in incidence, risk factors, stage at diagnosis, and survival. Established risk factors include tobacco use, alcohol consumption, areca nut use, and oncogenic human papillomavirus (HPV), although their relative contributions vary considerably across populations and regions.

Large international collaborative efforts have advanced understanding of HNC etiology by pooling epidemiologic data across studies.[2–4] However, these initiatives have primarily focused on risk factor identification and disease susceptibility. Comparatively less attention has been given to understanding differences in clinical presentation, care pathways, diagnostic and treatment pathways, survival, the role of HPV, and the molecular determinants of outcomes across diverse populations. Furthermore, many existing resources lack integration of detailed epidemiologic, clinical, follow-up, and biospecimen data within a common framework.

The HEADSpAcE (Head and Neck Cancer Epidemiology and Genomic Studies on Prognosis and Survival) Data Center was established to address these gaps by assembling a large, multi-country resource with harmonized epidemiologic, clinical, and biospecimen data across diverse populations. The Data Center currently includes data from over 26,000 participants (19,800 HNC cases and 7,600 controls, ie no cancer individuals) across 34 centers from 18 countries, encompassing a wide range of geographic regions, exposure profiles, and healthcare systems. Data are contributed through multiple studies and consortia, and harmonized within the HEADSpAcE Data Center under a shared governance framework agreed upon by participating principal investigators to support future collaborative research.

Specifically, the HEADSpAcE Data Center combines:

- sociodemographic data
- detailed exposure data
- standardized clinical data including treatment, and survival outcomes
- biospecimens supporting genomic and molecular analyses

The Data Center is intended to support a broad range of studies spanning cancer etiology, prevention, early detection, clinical outcomes, molecular characterization, while enabling research across diverse populations and healthcare settings.

Contributing Consortia and Studies

The HEADSpAcE Data Center integrates data and biospecimens from multiple HNC studies across North America, South America, Europe, and the Eastern Mediterranean between 1991–2025 (Table 1, Figure 1). These studies vary in design, recruitment strategy, follow-up procedures, and biospecimen availability, but collectively provide large-scale epidemiological, clinical, molecular, and longitudinal data spanning diverse populations, healthcare systems, and time periods. Contributing studies can be broadly grouped into IARC/WHO-coordinated studies and independent HNC-specific studies.

IARC/WHO-coordinated studies include ARCAGE,[5] InterCHANGE,[6] and the HEADSpAcE[7] cohort (Table 1). ARCAGE was originally established as a European multicentre case-control study of HNC risk factors and was subsequently expanded to incorporate clinical characteristics, HPV-related data, and patient follow-up. InterCHANGE established a large multicentre resource across South America. The HEADSpAcE cohort introduced a standardized protocol across centres in Europe and South America, including harmonized biospecimen collection and detailed assessment of diagnostic pathways and time-to-diagnosis intervals. These studies included biospecimen collections, which are primarily stored at IARC, and with matched and harmonized data integrated into the Data Center.

Additional independent studies contributing to the Data Center include HN5000[8,9] (United Kingdom), IROPICAN[10] (Iran), ICO Barcelona[11] (Spain), CHANCE[12] (USA), Pittsburgh[13] (USA), Toronto MSH-PMH[14,15] (Canada), and HN-iCTI[16,17] (Canada). Detailed descriptions of their design, recruitment, and data collection procedures have been reported elsewhere. Biospecimens from these independent studies remain at the contributing institutions and can be accessed through direct collaboration with the respective principal investigators, while harmonized data are incorporated into the Data Center framework.

Table 1. Studies Contributing to the Headspace Data Center.

Study	Region	#centres	Design	Recruitment Period	Number of Participants	Controls/ Cases	OPC	OC	LX	HPC	OTHER[#]	Biospecimen Location
ARCAGE[5]	Europe	12	Case-control	2002–2005	3966	2043/1923	426	560	652	163	122	IARC
INTERCHANGE[6]	South America	8	Case-Control	2011–2017	4235	1746/2489	681	1013	598	159	38	IARC
CHANCE[12]	USA	Many*	Case-control	2002–2006	2199	887/1312	383	366	485	31	47	Local site
U. PITTSBURGH[13]	USA	1	Case-control	2004–2014	2812	1050/1762	592	728	375	59	8	Local site
TORONTO MSH-PMH[14,15]	Canada	1	Case-control	2002–2018	5192	1894/3298	1353	909	508	91	437	Local site
HEADSPACE Cohort	Multi-region	17	Cohort	2019–2025	1650	0/1650	493	709	261	53	134	IARC
HN5000[8,9]	UK	76	Cohort	2011–2014	4694	0/4694	1914	1346	1062	238	134	Local site
IROPICAN[6]	Iran	Many^	Case-control	2011–2024	617	0/617	6	199	344	13	55	Local site
ICO[11]	Spain	1	Cohort	1991–2016	864	0/864	864	0	0	0	0	Local site
HN-iCTI[16,17]	Canada	1	Cohort	2010–2020	253	0/253	110	71	59	10	3	Local site

All studies include harmonized epidemiological, clinical, and follow-up data; availability of biospecimens and molecular data varies by study and is described in the data types section.
Abbreviations: OC, oral cavity cancer; OPC, oropharyngeal cancer; LX, laryngeal cancer; HPC, hypopharyngeal cancer.
*North Carolina cancer registry;
^across 10 provinces;
[#]Other includes unknown primary, Overlapping sites, pharynx, nasopharynx (HN5000 only) and missing subsite information.
PIs can be contacted directly for samples stored at their local site.

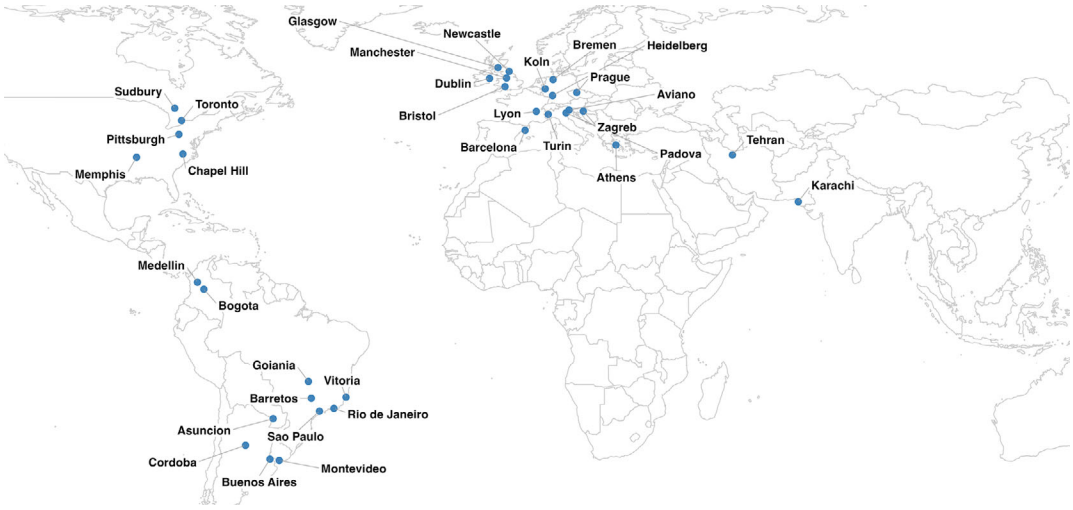


Figure 1. Map of the 34 centers across 18 countries currently comprising the HEADSpAcE Data Center. Base map data were obtained from the *maps* R package; study-centre locations were geocoded using OpenStreetMap. The figure was generated in R.

Collectively, all studies contribute epidemiological, clinical, molecular, and follow-up data from diverse populations and healthcare settings. Regardless of study design, cases were followed for at least five years after diagnosis to support survival and outcome analyses.

Building on these individual studies, three major IARC/WHO-led consortia—HEADSpAcE, VOYAGER,[18] and the GWAS initiative, were established. These consortia harmonized data across participating studies and generated extensive molecular, genomic, and biomarker datasets.

The HEADSpAcE consortium (Translational studies of HEAD and neck cancer in South America and Europe) operated from 2019 to 2025 and was funded by the European Commission (Horizon 2020; grant agreement No. 825771), Sao Paulo Research Foundation (FAPESP), and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). It brought together ARCAGE, InterCHANGE, HN5000, IROPICAN, and ICO Barcelona, and established a new patient cohort across South America and Europe, the HEADSpAcE cohort. This prospective recruitment established a contemporary cohort with detailed information on patient and healthcare system time intervals. The consortium's overall primary objective was to investigate determinants of HNC diagnosis and survival across diverse populations, focusing on delayed diagnosis, HPV-related disease, and genetic susceptibility. To support these aims, the consortium generated HPV biomarker and genomic sequencing data and healthcare systems data across recruitment centers, generating several important outputs. Data from the HEADSpAcE Consortium have shown increases in advanced-stage presentation during the COVID-19 pandemic, with regional heterogeneity,[19–22] and identified disparities in treatment guidelines that informed consensus recommendations in Latin America.[23,24] Additional work has characterized global incidence patterns, risk factors, and diagnostic pathways, including factors associated with time intervals of diagnosis.[25–27]

The VOYAGER consortium operated from 2017 to 2023, and was funded by the US National Institude of Dental and Craniofacial Health (R01DE025712). It integrated ARCAGE, CHANCE, Pittsburgh, Toronto, and HN5000 to examine genetic and environmental determinants of HNC risk and survival. Its primary objective was to investigate the genetic and environmental determinants of HNC risk and survival through the integration of epidemiological, clinical, and molecular data across multiple studies. The consortium harmonized genetic and epidemiological data across participating studies and generated large-scale genomic resources, enabling analyses of HNC susceptibility, HPV-related disease, and molecular tumour characteristics. The consortium has generated several major results, including the identification of novel susceptibility loci,[28] characterization of genetic risk across HNC subsites,[29,30] and evaluation of risk prediction models.[26,31] Detailed descriptions of VOYAGER data harmonization, structure, and published findings have previously been reported and are provided elsewhere.[18]

Finally, the Genome-Wide Association Study (GWAS) of oral cavity, pharynx, and larynx cancers was funded by the US National Cancer Institute and US National Institude of Dental and Craniofacial Health (R03DE030257, X01HG010743). Building on earlier collaborative GWAS efforts, including the OncoArray study,[2,4,32] this initiative brought together all

studies from the HEADSpAcE and VOYAGER consortia, and additionally included HN-iCTI. This effort resulted in the largest multi-ancestry HNC GWAS to date, identifying 29 susceptibility loci.[33]

To support coordinated access and analysis of these resources, IARC/WHO established the HEADSpAcE Data Center as a common framework that brings together data and biospecimens generated through IARC/WHO-led HNC studies and partner networks (Table 1). The Data Center provides an infrastructure for standardized integration, governance, and access across existing and future studies. This resource currently includes harmonized epidemiological, clinical, demographic, follow-up, and sequencing data from over 27,000 individuals across 18 countries, representing diverse populations and healthcare settings.

By facilitating harmonized access and analysis across diverse datasets, the Data Center supports large-scale analyses of HNC risk, diagnosis, and survival, and serves as a platform for global collaboration to address persistent gaps in cancer epidemiology, early detection, and prognosis.

Data Center Resources

The HEADSpAcE Data Center currently includes over 27,000 participants (~18,800 oral cavity, oropharynx, larynx and hypopharynx cancer cases, ~1000 other HNC sites, and ~7600 controls) with harmonized clinical, epidemiological, and lifestyle data.

Cases are defined using ICD-10 codes as follows: oral cavity (C00.3-C00.6, C00.8-C00.9, C02.0-C02.3, C02.8, C02.9, C03, C04, C05.0, C05.8, C05.9, C06), oropharynx (C01, C02.4, C05.1-C05.2, C09.0-C10.9), larynx (C32), hypopharynx (C12, C13), and Others (unknown primary, overlapping sites, pharynx, nasopharynx (HN5000 only) and missing subsite).

Currently, 125 variables across studies have been harmonized, including 1) demographic details (age, sex, country, ethnicity, education); 2) anthropometric data (height, weight at diagnosis and prior to diagnosis, body mass index); 3) lifestyle exposures (smoked tobacco use, alcohol consumption, opium use (Iran), smokeless tobacco use (Pakistan), oral health/hygiene, sexual behaviors partners, consumption of hot drinks); 4) clinical characteristics (ICD code and tumor site, TNM staging, diagnostic and other HPV-related biomarkers, treatment type, treatment duration, recurrence, metastasis, second primary cancers, follow-up time, vital status).

All data are de-identified locally using a standardized alphanumeric identifier before transfer to the HEADSpAcE Data Center. This harmonized structure enables consistent cross-study analyses across populations and study designs.

Tumor stage was harmonized according to the AJCC 8th edition,[34] which was the applicable edition at the time of harmonization. Reclassification was restricted to p16-positive oropharyngeal cancers because these were the tumors for which the available staging information permitted informative conversion from the 7th to the 8th edition. Clinical and pathological stage groups were reassigned using available T, N, and M information. For pathological staging, the number of positive lymph nodes required under the 8th edition was not consistently available. Pathological stage was therefore approximated using the available 7th-edition pathological N categories. Among the 22 p16-positive oropharyngeal cancers requiring pathological stage conversion, 11 (50%) could not be mapped exactly, representing a very small proportion of the overall harmonized dataset (<0.1%).

Clinical, Demographic, and Follow-up Profile

Tumor subsites include oral cavity (32%), oropharynx (39%), larynx (19%), hypopharynx (4%), and other head and neck sites (6%) (Figure 2).

Across all subsites, the majority of patients are male, with the highest proportion observed in laryngeal cancer (88%). Body mass index distributions at diagnosis were broadly similar across subsites, with most patients classified as normal weight or overweight.[35] Median age at diagnosis ranged from 59 years in oropharyngeal cancer to 63 years in laryngeal cancer. The majority of patients presented with advanced-stage disease.

Exposure profiles varied by subsite. Smoking and alcohol use are most prevalent in laryngeal and hypopharyngeal cancers, while oropharyngeal cancers showed a more heterogeneous exposure profile. HPV-related biomarkers (mainly serology or p16) were available primarily for oropharyngeal tumors.

Follow-up data are available for the majority of patients in the HEADSpAcE Data Center (Figure 3). Overall survival information is currently available for 95% of patients, while recurrence and information on second primary tumour are

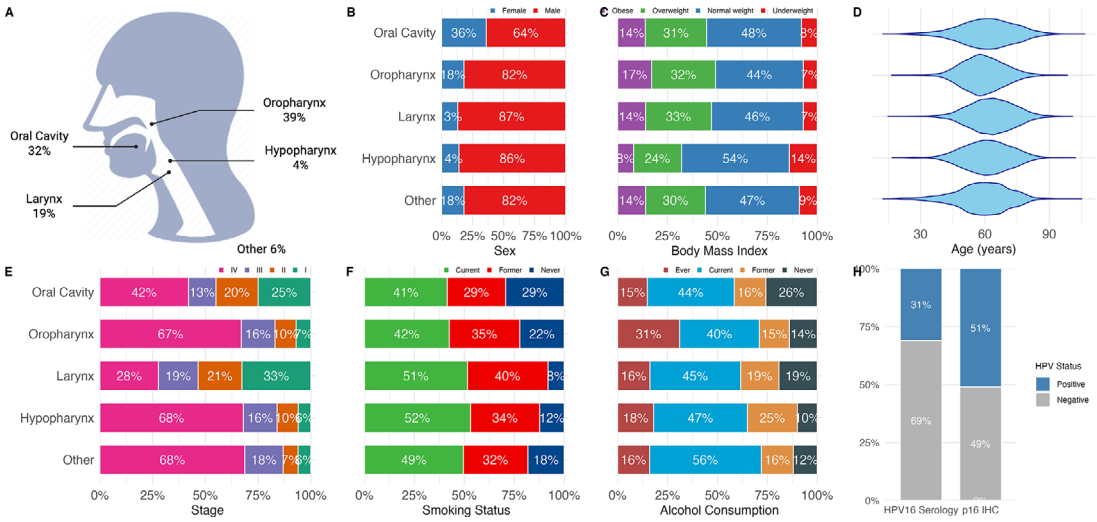


Figure 2. Data Center Snapshot. A) proportions of each subsite; Distribution of cases across subsites by B) Sex; C) BMI at diagnosis; D) Age at diagnosis; E) Tumor Stage; F) Smoking Status; G) Alcohol Consumption and F) HPV status, available by either p16 IHC or HPV16 serology primarily, in oropharynx tumors. For 60% of oropharyngeal cancer patients, both markers are available.

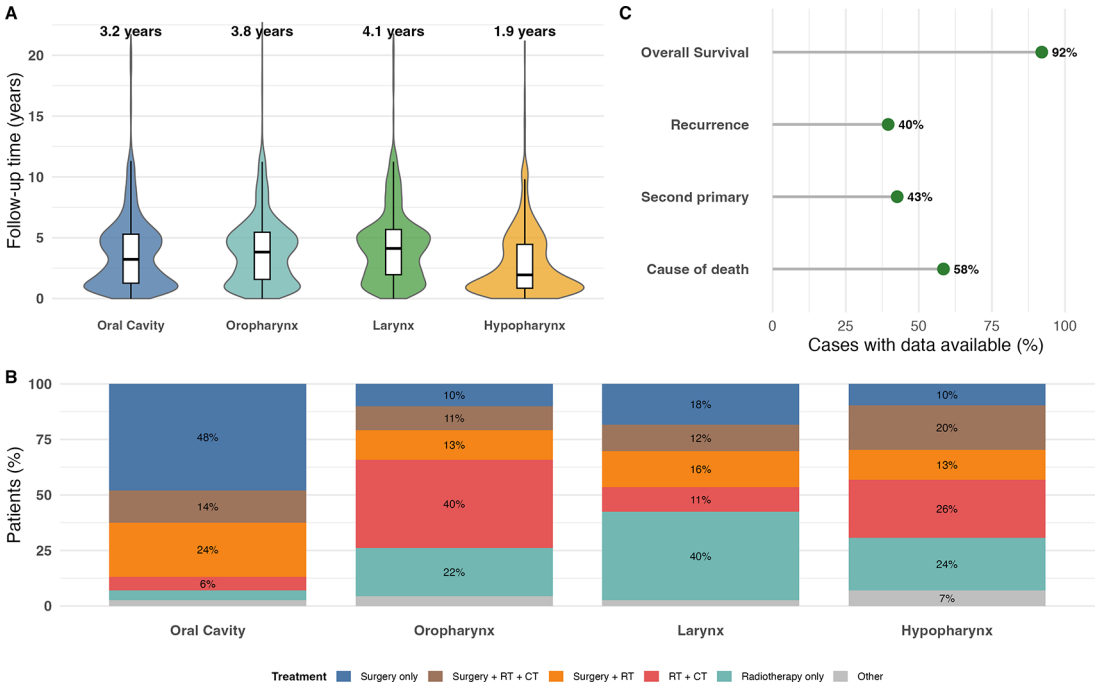


Figure 3. Treatment patterns and follow-up data in the HEADSpAcE Data Center. (A) Distribution of overall follow-up time by tumor subsite. Labels indicate the median follow-up time. (B) Distribution of primary treatment modalities among patients with complete treatment information, stratified by tumor subsite. "Other" includes chemotherapy only and patients who did not receive surgery, radiotherapy, or chemotherapy. (C) Availability of key outcome variables within the harmonized dataset. Percentages represent the proportion of patients with available data for overall survival, recurrence, and second primary tumors. Cause-of-death availability was calculated among deceased patients only.

available for 47% and 48% of patients, respectively. Cause of death was recorded for 58% of deceased patients. Median follow-up times ranges from 2.0 years for hypopharynx cancers to 4.2 years for larynx cancers. Treatment patterns vary across tumor subsites, reflecting established standards of care. Surgery alone is the predominant treatment for oral cavity

cancers (47%), whereas combined radiotherapy and chemotherapy are the most common treatment for oropharyngeal cancers (42%). Laryngeal cancers are most frequently treated with radiotherapy alone (33%), while hypopharyngeal cancers show the greatest use of multimodality treatment, including radiotherapy with chemotherapy (27%) and trimodality therapy comprising surgery, radiotherapy, and chemotherapy (18%) (Figure 3).

The Data Center includes substantial representation of key subgroups, including never smokers (3047 cases, 16%), never drinkers (2373 cases, 13%), younger individuals (<45 years; 1298 cases, 7%), and patients with less common or clinically aggressive disease, such as hypoharynx cancer (745 cases, 4%). By combining data across multiple international studies, the Data Center provides adequate sample sizes to investigate rare tumor subsites and patient populations that are generally too small to study within individual cohorts. This resource therefore enables analyses of subgroup-specific risk factors, molecular characteristics, treatment patterns, and clinical outcomes while maintaining sufficient statistical power.

Diagnostic Pathway and Time Intervals Data (HEADSpAcE Cohort)

The HEADSpAcE cohort, which was the prospective recruitment from the European-Commission funded consortium, includes 1,650 patients diagnosed with cancers of the oral cavity, oropharynx, hypopharynx, or larynx and recruited through 17 centres across Europe, Latin America, and Asia. The cohort includes detailed demographic, lifestyle, clinical, treatment, and annual follow-up data, together with linked biospecimens. Depending on the centre, specimens collected at recruitment include whole blood, plasma, serum, tumour tissue, adjacent normal tissue, oral rinse samples, and buccal swabs. Some centres also collect blood, oral rinse samples, or buccal swabs during annual follow-up, providing longitudinal biospecimens for evaluating changes after diagnosis and treatment. The cohort additionally captures patient- and health system–level factors associated with diagnostic time intervals that are not available in the other contributing datasets (Figure 4). The study questionnaire captures psychosocial, financial, structural, and educational barriers to timely diagnosis across the diagnostic pathway, from symptom onset and interpretation to the first healthcare contact and subsequent progression through the healthcare system. Additional information includes travel time, healthcare access, and self-reported reasons for delay. The study protocol and questionnaire are available on the cohort website (https://headspace.iarc.who.int/contributing-studies/).

These data enable reconstruction of patient trajectories from symptom onset through diagnosis within a standardized framework adapted from the Aarhus Statement.[36] This component provides a unique resource to evaluate diagnostic pathway time intervals across diverse healthcare settings. A center-level benchmarking analysis of health system factors associated with advanced-stage diagnosis has been reported previously[27] and complementary qualitative analyses of center case-studies from the HEADSpAcE consortium have identified key system-level barriers to timely diagnosis,

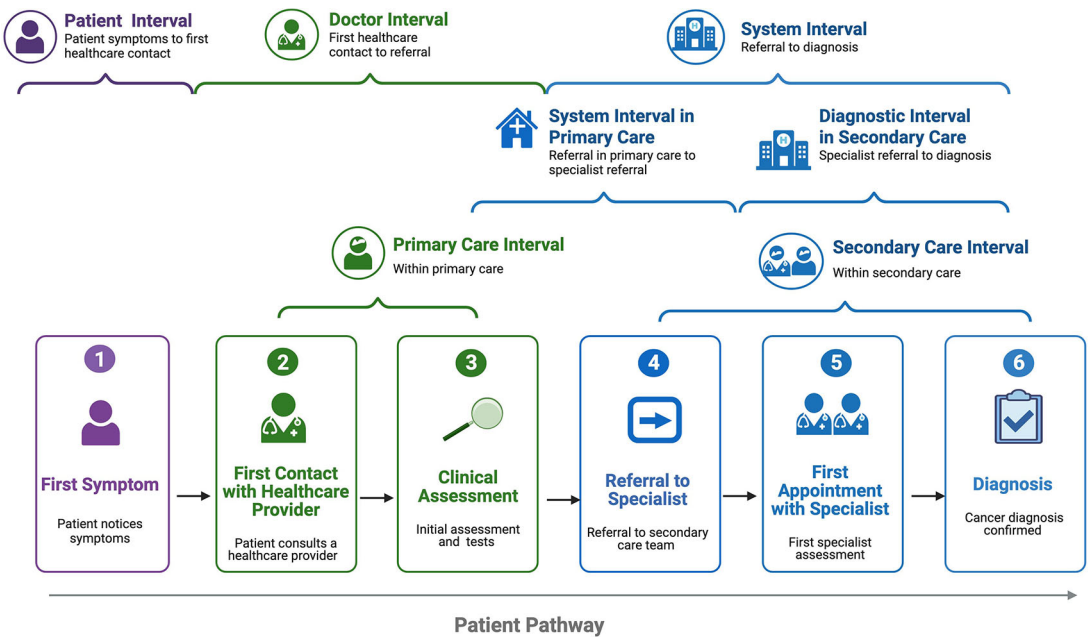


Figure 4. Data collected to reconstruct patient diagnostic pathways. Created in https://BioRender.com

including access to care, navigation of diagnostic pathways, socioeconomic and geographic inequalities, and coordination between healthcare teams.[37]

Molecular Data

The HEADSpAcE Data Center centralizes sequencing and assay data generated by various initiatives and consortia. These include immunohistochemical, serological, in situ hybridization, and high-throughput sequencing assays, that were implemented to address key biological questions in HNC progression and prognosis.

Data from six molecular assays are currently available (Table 2), with prioritization of samples guided by tissue quality, anatomical site, plasma availability, and the scientific objectives of the originating studies.

Omics data from six assay types are currently available (Table 3), comprising targeted DNA sequencing, DNA methylation profiling, SNP genotyping, whole-genome sequencing, whole-exome sequencing, and targeted RNA expression. Sample selection for each assay was guided by biospecimen availability and quality, anatomical site, and the scientific objectives of the originating studies.

Table 2. Biomarker assays available in the HEADSpAcE Data Centre.

Biomarker	N	Sample type	Method	Key details
HPV16 serology	9 991	Serum	Luminex multiplex assay[43,44]	Antibodies to HPV16 viral proteins
p16 immunohistochemistry	7 264	FFPE tumour tissue	Immunohistochemistry	Positivity defined as moderate-to-strong, diffuse nuclear and cytoplasmic staining in ≥70% of tumour cells[45]
HPV DNA	1 064	FFPE tumour tissue	PCR or in situ hybridization	Detection of common high-risk HPV genotypes
HPV E6/E7 RNA	470	FFPE tumour tissue	RNAscope v2.0 (Advanced Cell Diagnostics, Hayward, CA, USA)	Target-specific E6 and E7 probes for HPV16, HPV18, HPV31, HPV33, HPV35, HPV52, and HPV58
PD-L1 expression	193	FFPE tumour tissue	22C3 assay on a Dako Autostainer (Agilent Technologies, CA, USA)	Quantification of programmed death-ligand 1 expression

Abbreviations: FFPE, formalin-fixed paraffin-embedded; HPV, human papillomavirus; IHC, immunohistochemistry; ISH, in situ hybridization; PCR, polymerase chain reaction; PD-L1, programmed death-ligand 1. N indicates the number with available assay data.

Table 3. Omics data available in the HEADSpAcE Data Centre.

Data type	N	Platform or kit	Key specifications
Targeted DNA sequencing	1 389	Custom ~1,100-gene panel[41]; Agilent SureSelect	Hybrid-capture sequencing; mean coverage, 500×
DNA methylation	482	Illumina Infinium MethylationEPIC BeadChip (850 K)	Bisulfite-converted tumour DNA
SNP genotyping	18 657	Illumina OncoArray, Global Screening Array, or All of Us array	Germline variation
Whole-genome sequencing	249	Illumina NovaSeq 6000	150-bp paired-end sequencing; mean depth, 40× tumour and 20× normal
Whole-exome sequencing	298	Agilent SureSelect Human All Exon V7 (part no. 5191–4005; Agilent, Santa Clara, CA, USA)	~35.7-Mb target; GRCh38
RNA expression	112	NanoString nCounter (NanoString, Seattle, WA, USA)	164-gene panel; FFPE-compatible

Abbreviations: FFPE, formalin-fixed paraffin-embedded; SNP, single-nucleotide polymorphism. N indicates the number with available assay data.

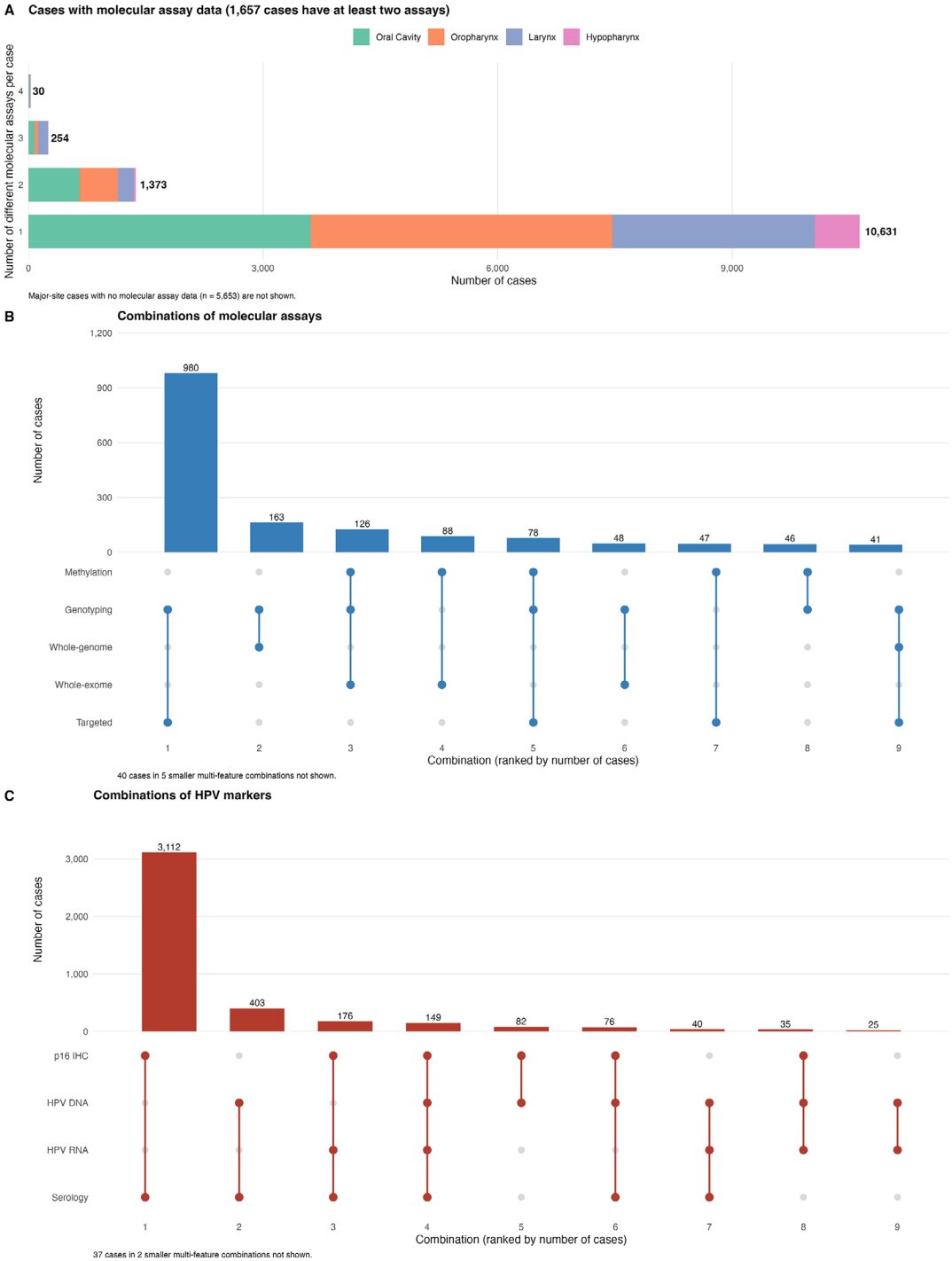


Figure 5. Molecular assay and HPV marker availability and overlap among head and neck cancer cases. (A) Cases with molecular assay data. Number of cases with one to four different molecular assays, stacked by tumour subsite. Molecular assays included methylation, genotyping, whole-genome sequencing, whole-exome sequencing and targeted sequencing. **(B)** Combinations of molecular assays. UpSet plot showing the nine most frequent combinations among cases with at least two molecular assays. **(C)** Combinations of HPV markers. UpSet plot showing the nine most frequent combinations among cases with at least two HPV markers: p16 immunohistochemistry (IHC), HPV DNA, HPV RNA and HPV serology. In panels B and C, vertical bars indicate the number of cases in each combination, and connected dots identify the assays or markers included. Assay and marker indicators represent data availability, not positive test results.

To illustrate the breadth and integration of molecular profiling within the HEADSpAcE Data Center, Figure 5 summarizes the overlap of molecular assays available for the same participants or tumour samples. Overlapping molecular data were available across the major HNC subsites, with oral cavity and oropharyngeal cancers accounting for the largest numbers (Figure 5A). The most common omics profile combined SNP genotyping with targeted DNA sequencing, while smaller subsets had other combinations of genotyping, targeted sequencing, whole-exome or whole-genome sequencing, and DNA methylation data (Figure 5B). Among HPV-related biomarkers, the most common combination comprised p16 immunohistochemistry, HPV DNA testing, and HPV16 serology, while smaller subsets had other combinations that also included HPV RNA testing (Figure 5C).

The availability of complementary molecular data—including germline and tumour sequence variation, DNA methylation, copy-number alterations, RNA expression, and HPV-related biomarkers—supports integrative analyses of HNC development, progression, and prognosis.

Sample Biorepository

The HEADSpAcE biorepository located at IARC/WHO integrates biospecimens collected across multiple studies, recruitment sites, and timepoints. Although sample collection followed study-specific protocols, all specimens are centrally catalogued using a harmonized metadata framework, enabling linkage to clinical, molecular, and follow-up data within the Data Center.

Biospecimens were collected at baseline and, whenever possible, during clinical follow-up. visits The most widely available sample types across regions are blood-derived specimens (plasma and buffy coat) from casesa and controls, with highest coverage in South America and Southern Europe. FFPE tumor tissue is available across all regions and represents the most consistently collected tumor material, reflecting routine clinical practice. Fresh frozen tumor and matched normal tissues are available for a subset of patients, with limited availability overall. Additional biospecimens include whole blood, serum, and, in some centers, saliva or oral rinse samples. Follow-up collections primarily consist of blood-based or saliva-derived specimens.

The distribution of biospecimens varies across regions, reflecting differences in clinical workflows, infrastructure, and study-specific collection protocols (Table 4). Standard operating procedures for sample collection for studies contributing to the IARC/WHO biorepository are publicly available on the HEADSpAcE website.[38]

Data Availability, Access, Governance and Expansion

The HEADSpAcE Data Center is an open and expanding resource that enables harmonized access to data and biospecimens while allowing contributing institutions to retain ownership of their data/samples. Each member institution is represented by a designated local Principal Investigator (PI), who serves as the liaison with IARC/WHO. Any use of contributed data or biospecimens requires approval from the relevant center PI and must comply with the HEADSpAcE Data Access Policy and Publication Guidelines, which are available at www.headspace.iarc.who.int.

Table 4. Availability of Biospecimens by Subregion.

	South America (4903 patients)	Northern Europe (5541 patients)	Western Europe (1297 patients)	Eastern Europe (457 patients)	Southern Europe (3096 patients)	Eastern Mediterranean* (732 patients)
Plasma	70% (3435)	2% (131)	29% (371)	94% (428)	46% (1416)	7% (53)
Buffy coat	57% (2784)	2% (131)	10% (134)	31% (141)	42% (1291)	7% (53)
Buccal Cells	9% (453)	2% (108)	0% (0)	0% (0)	10% (325)	6% (47)
Oral Rinse	11% (517)	0% (0)	0% (0)	0% (0)	0% (0)	2% (14)
Fresh/Frozen Tumor	18% (862)	0% (0)	8% (106)	31% (143)	7% (218)	7% (49)
Fresh/Frozen Normal	13% (660)	0% (0)	8% (106)	26% (118)	0% (0)	7% (52)
FFPE Tumor	15% (724)	4% (231)	24% (315)	48% (221)	17% (525)	18% (134)
FFPE Normal	0% (4)	0% (0)	0% (0)	0% (0)	1% (30)	4% (27)

Numbers indicate biospecimens stored at the IARC/WHO biorepository and governed under the HEADSpAcE Data Center.
*Includes Iran and Pakistan. The percentage refers to % within a subregion. Samples are reported irrespective of case/control status.

Researchers seeking access can submit a proposal through the HEADSpAcE website, which must be in collaboration with a current and active HEADSpAcE Data Center Member. Applications are reviewed quarterly by the HEADSpAcE Data Use Committee based on scientific quality, alignment with the Data Center's scientific mandate, ethical compliance, and potential overlap with ongoing projects. Supporting resources, including descriptions of contributing cohorts, patient recruitment protocols, references to primary study publications, and a harmonized data dictionary of available variables, are available through the HEADSpAcE Data Centerwebsite. To date, 19 research proposals have been approved, spanning descriptive and molecular epidemiology, HPV-related biomarkers, genetic and tumour determinants, prognostic modelling, socioeconomic inequalities, early diagnosis, and time to care.

The governance framework supports continued expansion of the Data Center through the inclusion of new participating institutions contributing data and/or biospecimens under agreements with IARC/WHO. New centers, particularly those from low- and middle-income countries and other underrepresented regions, are actively encouraged to participate, facilitating international collaborative research while improving the geographic diversity of the resource. Although the current Data Center includes studies from Europe, North and South America, and the Middle East, representation from Southeast Asia, Africa, and the Western Pacific remains limited despite these regions bearing a substantial proportion (63%) of the global HNC burden.[1] Expanding participation from these regions is a priority to improve the global representativeness of the resource and enable more comprehensive investigations of regional differences in exposure patterns, HPV prevalence, tumor biology, clinical management and outcomes.

Several molecular datasets are also available through controlled-access public repositories. Germline genotyping data are available through dbGaP under accession numbers phs001202.v2.p1 and phs002503.v1.p1, with somatic sequencing data currently being deposited. Whole-genome sequencing data are available through the European Genome-Phenome Archive (EGA; study EGAS00001005450) and the ICGC-ARGO platform. Within the EGA, aligned BAM files are available under dataset EGAD00001015386, consensus SNV and indel variant calls under EGAD00001015388, patient metadata under EGAD00001015387, structural variant calls under EGAD00001015389, and copy number variant calls under EGAD00001015390. While these repositories provide access to genomic data, the accompanying participant phenotypic information is is necessarily limited. The HEADSpAcE Data Center links these molecular datasets to rich harmonized clinical, epidemiological, treatment, and outcome data, enabling more comprehensive integrative analyses.

Ethics approval

The overall project was approved by the IARC Ethics Committee (19–23-A7). Once a center obtained local Institutional Review Board approval, it was included under the Data Center ethics approval.

Disclaimer

TW serves on advisory boards for MSD (Merck) Sharp & Dohme. GL serves on advisory boards of Amgen, AstraZeneca, AdelaBio, Bayer, Boehringer Ingelheim, Bristol Myers Squibb, Daiichi Sankyo, Dizal Pharma, EMD Serono, Gilead, GlaxoSmithKline, Jazz Pharmaceuticals, Johnson and Johnson, Lilly, Merck, Novartis, Nuvation Bio, Nuvalent, OxCan, Pfizer, Regeneron, Roche, and Takeda. PS reports research grant support from ADMIR, HTG Molecular Diagnostics, Bristol Myers Squibb, Roche, AstraZeneca, Novartis, BMS Foundation and Illumina.

Where members are identified as personnel of the International Agency for Research on Cancer/World Health Organization, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer/World Health Organization.

HEADSpAcE Consortium Members

Béatrice Fervers[1], Roque Adam[2], Laurie Ailles[3], Salima Alibhai[4], Miquel Angel Pavon[5], Namrah Anwar[6], Paola Engelmann Arantes[7], Julia Ferreira Araújo[8], Alanis Arévalo[9], Lisa Arguello[10], Lucas Avondet[2], Camila Batista Daniel[11], Paola Engelmann Alejandra Belardinelli[12], Bianca Beraldi[13], Barbara Berenstein[14], Natália Bernardino Rodrigues[15], Jaroslav Betka[16], Marianna G Botta[7], Scott Bratman[17], Nicole Brenner[18], Carol Brentisci[19], María Luisa Cabañas[20], Raiany S Carvalho[21], Andre L Carvalho[22], Luigi Chiusa[23], Priscilla Chopard[24], Qurratulain Chundriger[25], Sheila Coelho Soares Lima[26], Isabela Costa[26], Marcelo Lucas Rodrigues Costalonga[8], Grant Creaney[27], Tauana C Dias[21], Evandro Duccini de Souza[13], Lais Corsino Durant[28], Ebrahimi Elmira[24], Alberto Escallón[29], Gisele Aparecida Fernandes[7], Diego Rodolfo Fernandez[30], Valentina Fiano[31], Elisabetta Fratta[32], Valerie Gaborieau[24], Ricardo R Gama[22], Manuela Gangemi[19], Paolo Garzino-Demo[33], Raul Giglio[14], Mariel A Goulart[27], Jéssica Graça Sant'Anna[11], Marek Grega[34], Anna Clara Gregório Có[11], Jose Antonio Hakim[29], Rashidian Hamideh[35], Marco Homero de Sá Santos[13], Katrina Hurley[36], Magalí Insfran[9], Giuseppe Carlo Iorio[37], Moghira Iqbaluddin Siddiqui[38], Jannik Johannsen[39], Martin Kaňa[40], Luiz Paulo Kowalski[41], Evelio Legal[42], Jeferson Lenzi[13], Fernando Luiz Dias[26], Iván Lyra González[43], Gary Macfarlane[44], Willene Machado Zorzaneli[11], Stephanie MacNeil[45], Ricardo Mai Rocha[13], Priscila Marinho de Abreu[11], Hadji Maryam[46],

James McCaul[47], James McKay[24], Alex D McMahon[27], Carlos Mena[42], Elismauro Francisco Mendonça[48], Lorena Meza[9], Birgitta Michels[18], Matinair Siqueira Mineiro[48], Chiara Moccia[31], Pamela Mongelos[9], Ana Lorena Montealegre-Páez[49], Rosana Andrea Andrea Morelatto[12], Francisca Morey Cortes[5], Jessica Paola Mosmann[50], Andy Ness[36], Aline B Neves[7], Marc Oliva[5], Hernán Ortiz[10], José Ortiz[42], Marta Osorio[42], Oliviero Ostellino[51], Claire Paterson[52], Sonia Paytubi Casabona[5], Giancarlo Pecorari[33], David M Pereira[53], Olivia Pérol[1], Alicia Pomata[20], Maja Popovic[31], Alisson Poveda[54], Carol P Prado[7], Kristina M Prager[18], Guglielmo Ramieri[33], Saida Rasul[4], Juliana Nunes Iglesias Rego[55], Rui M Reis[21], Helene Renard[24], Katrina Rey-McIntyre[17], Umberto Ricardi[37], Giuseppe Riva[33], Frederic Rodilla[5], Ingrid Rodriguez[56], María Isabel Rodríguez[9], Alastair Ross[27], Tazeen Saeed Ali[57], Pierre Saintigny[58], Cristóvam Scapultampo-Neto[59], Agenor Sena[13], Mokhtari Sepideh[35], Ricardo Serrano[42], Shachi Jenny Sharma[39], Oliver Siefer[39], Bruna Pereira Sorroche[21], Cinthia Sosa[20], Juliana de Oliveira Souza[28], Antonella Stura[19], Steven Thomas[36], Stu Toms[36], Sara Tous[5], Adriana Valenzuela[9], Isabella Bittencourt Valle[11], Alex Whitmarsh[36], Sylvia Wright[60].

[1]Department Cancer Environnement, Centre Léon Bérard, France.

[2]H&N cancer Department, Universidad de Buenos Aires, Argentina.

[3]Medical Biophysics, Princess Margaret Cancer Centre, University of Toronto, Canada.

[4]Department of Surgery, Dental Hygiene Program, Aga Khan University Hospital, Pakistan.

[5]Catalan Institute of Oncology (ICO), Spain.

[6]Faculty of Science and Technology, University of Central Punjab, Pakistan.

[7]Group of Epidemiology and Statistics on Cancer, A.C. Camargo Cancer Center, Brazil.

[8]Postgraduate Program in Biotechnology, Associação Feminina de Educação e Combateao Câncer (AFECC), Universidade Federal do Espirito Santo, Brazil.

[9]Salud Pública, Instituto de Investigaiones en Ciencias de la Salud (IICS), Universidad Nacional de Asunción (UNA), Paraguay.

[10]Servicio de Cabeza y Cuello, Instituto Nacional del Cáncer, Ministerio de Salud Pública y Bienestar Social, Paraguay.

[11]Postgraduate Program in Biotechnology, Universidade Federal do Espirito Santo, Brazil.

[12]Oral Pathology, School of Dentistry, National University of Cordoba (AR), Dpt. Oral Pathology, Argentina.

[13]Head and Neck Surgery Division, Associação Feminina de Educação e Combateao Câncer (AFECC), Hospital Santa Rita de Cássia, Brazil.

[14]H&N cancer Department, Institute of Oncology Angel H. Roffo, University of Buenos Aires, Argentina.

[15]Department of Pathology, Universidade Federal do Espirito Santo, Brazil.

[16]Departrment of Otorhinolaryngology and Head and Neck Surgery, 1st Faculty of Medicine, University Hospital in Motol and Charles University, Czech Republic.

[17]Radiation Oncology Department, Princess Margaret Cancer Centre, University of Toronto, Canada.

[18]Division of Infections and Cancer Epidemiology, German Cancer Research Center (DKFZ), Germany.

[19]Department of Medical Sciences, Cancer Epidemiology Unit, AOU Città della Salute e della Scienza di Torino, Italy.

[20]Departamento de Anatomía Patológica, Instituto Nacional del Cáncer, Ministerio de Salud Pública y Bienestar Social, Paraguay.

[21]Molecular Oncology Research Center, Barretos Cancer Hospital, Brazil.

[22]Department of Head and Neck Surgery, Barretos Cancer Hospital, Brazil.

[23]Pathology Unit, AOU Città della Salute e della Scienza di Torino, Italy.

[24]Genomic Epidemiology Branch, International Agency for Research on cancer (IARC/WHO), France.

[25]Department of Pathology and Laboratory Medicine, Section of Histopathology, Aga Khan University Hospital, Pakistan.

[26]INCA, Brazil.

[27]School of Medicine, Dentistry and Nursing, University of Glasgow, United Kingdom.

[28]A.C. Camargo Cancer Center, Brazil.

[29]Department of Surgery, Head and Neck Division, Fundación SantaFe de Bogotá, Colombia.

[30]Centro de Radioterapia Dean Funes, Privated, Argentina.

[31]Department of Medical Sciences, Cancer Epidemiology Unit, University of Turin, Italy.

[32]Immunopathology and Cancer Biomarkers, Centro di Riferiemnto Oncologico di Avinao (CRO) IRCCS, Italy.

[33]Department of Surgical Sciences, University of Turin, Italy.

[34]Department of Pathology and Molecular Medicine, 2nd Faculty of Medicine, Charles University and Motol University Hospital, University Hospital in Motol, Czech Republic.

[35]Cancer Research Center, Tehran University of Medical Sciences, Iran.

[36]Bristol Dental School, University of Bristol, United Kingdom.

[37]Department of Oncology, University of Turin, Italy.

[38]Department of Surgery, Section of E.N.T, Aga Khan University Hospital, Pakistan.

[39]Department of Otorhinolaryngology, Head and Neck Surgery, University of Cologne, Germany.

[40]Departrment of Otorhinolaryngology and Head and Neck Surgery, University Hospital in Motol, Czech Republic.

[41]Head and Neck Surgery and Otorhinolaryngology Department, Oncology Surgery, A.C. Camargo Cancer Center, Brazil.

[42]Cátedra Otorrinonaringología, Hospital de Clínicas, Facultad de Ciencias Médicas, Universidad Nacional de Asunción, Paraguay.

[43]Servicio de Oncología Clínica Hospital de Clínicas, Universidad de la República, Uruguay.

[44]Epidemiology Group, University of Aberdeen, UK.

[45]Glasgow Dental Hospital and School, NHS Greater Glasgow and Clyde, United Kingdom.

[46]Faculty of Social Science, Tampere University, Finland.

[47]Department of Oral and Maxillofacial/Head and Neck Surgery, NHS Greater Glasgow and Clyde, United Kingdom.

[48]Hospital Câncer Araújo Jorge, Brazil.

[49]Faculty of Medicine, El Bosque University, Colombia.

[50]Instituto de Virología Dr.J.M.Vanella, Facultad de Ciencias Medicas, Enf. Gordillo Gomez s/n. Ciudad Universitaria, Argentina.

[51]Department of Oncology, Division of Medical Oncology, AOU Città della Salute e della Scienza di Torino, Italy.

[52]Beatson West of Scotland Cancer Centre, NHS Greater Glasgow and Clyde, United Kingdom.

[53]Radiation Oncology Department, Institute of Oncology Angel H. Roffo, University of Buenos Aires, Argentina.

[54]Faculy of Medicine, El Bosque University, Colombia.

[55]Clinical Research Center, Associação Feminina de Educação e Combateao Câncer (AFECC), Hospital Santa Rita de Cássia, Brazil.

[56]Laboratorio de Anatomía Patológica, Hospital de Clínicas, Facultad de Ciencias Médicas, Universidad Nacional de Asunción, Paraguay.

[57]School of Nursing and Midwifery, Aga Khan University Hospital, Pakistan.

[58]Centre Léon Bérard, France.

[59]Pathology and Molecular Diagnostics Service, Barretos Cancer Hospital, Brazil.

[60]Institute of Cancer Sciences, University of Glasgow, United Kingdom.

Data Availability

Underlying data

Zenodo: *HEADSpAcE Data Center: source data for "The HEADSpAcE Data Center: An International Data and Biospecimen Resource for Head and Neck Cancer Translational Research".* https://doi.org/10.5281/zenodo.22310158[39]

The project contains the following underlying data:

- HEADSpAcE_figure_source_data.xlsx (aggregate counts, percentages, and summary values underlying the figures in this article).

Creative Commons Attribution 4.0 International (CC BY 4.0).

dbGaP: *OncoArray: Oral and Pharynx Cancer.* phs001202.v2.p1.[40]

The project contains the following underlying data:

- Germline genotyping data.

These data are available to approved researchers through dbGaP controlled access, subject to its Data Use Certification and study-specific data use limitations.

dbGaP: *GWA Study of Oral Cavity, Pharynx and Larynx Cancers in European, and North and South American Populations – CIDR.* phs002503.v1.p1[41]

The project contains the following underlying data:

- Germline genotyping data.

These data are available to approved researchers through dbGaP controlled access, subject to its Data Use Certification and study-specific data use limitations.

European Genome-phenome Archive: *[study title].* EGAS00001005450[42]

The project contains the following underlying data:

- EGAD00001015386 (aligned BAM files).
- EGAD00001015387 (participant metadata).
- EGAD00001015388 (consensus SNV and indel variant calls).
- EGAD00001015389 (structural-variant calls).
- EGAD00001015390 (copy-number variant calls).

Individual-level HEADSpAcE clinical, epidemiological, treatment, outcome, imaging, and biospecimen-linked data, and the dbGaP and EGA datasets listed above, are controlled access because of consent, ethical, privacy, and contributing-centre data-sharing requirements. Access to HEADSpAcE data may be requested through https://headspace.iarc.who.int.[7]

Extended data

HEADSpAcE Data Center: *HEADSpAcE Data Center website.* https://headspace.iarc.who.int.[7]

This resource contains the following extended data and documentation:

- Descriptions of contributing cohorts and participating centres.
- Patient recruitment protocols and links to primary study publications.
- Harmonized data dictionary, including variable descriptions and coding.
- Information on available clinical, epidemiological, treatment, outcome, imaging, molecular, and biospecimen data.
- HEADSpAcE Data Access Policy and Publication Guidelines.
- Information on submitting data-access proposals and the governance framework for access to the Data Center.

Acknowledgments

We are grateful for the inclusion of the ARCAGE (Manchester centre) data by agreement with Professor Gary J Macfarlane, University of Aberdeen, UK.

The authors honor the memory of Dr. José Roberto V. de Podestá and gratefully acknowledge his outstanding leadership and dedication to head and neck cancer research at the Associação Feminina de Educação e Combate ao Câncer (AFECC)–Hospital Santa Rita de Cássia and to the HEADSpAcE Data Center.

References

1. Sung H, Filho AM, Laversanne M, *et al.*: **Global cancer statistics 2024: GLOBOCAN estimates of incidence and mortality worldwide for 34 cancers in 186 countries.** *CA Cancer J. Clin.* Jul-Aug 2026; **76**(4): e70090.
PubMed Abstract | Publisher Full Text | Free Full Text

2. Lesseur C, Ferreiro-Iglesias A, McKay JD, *et al.*: **Genome-wide association meta-analysis identifies pleiotropic risk loci for aerodigestive squamous cell cancers.** *PLoS Genet.* Mar 2021; **17**(3): e1009254.
PubMed Abstract | Publisher Full Text | Free Full Text

3. Bravi F, Lee YA, Hashibe M, *et al.*: **Lessons learned from the INHANCE consortium: An overview of recent results on head and neck cancer.** *Oral Dis.* Jan 2021; **27**(1): 73–93.
PubMed Abstract | Publisher Full Text | Free Full Text

4. Lesseur C, Diergaarde B, Olshan AF, *et al.*: **Genome-wide association analyses identify new susceptibility loci for oral cavity and pharyngeal cancer.** *Nat. Genet.* Dec 2016; **48**(12): 1544–1550.
PubMed Abstract | Publisher Full Text | Free Full Text

5. Lagiou P, Georgila C, Minaki P, *et al.*: **Alcohol-related cancers and genetic susceptibility in Europe: the ARCAGE project: study samples and data collection.** *Eur. J. Cancer Prev.* Feb 2009; **18**(1): 76–84.
PubMed Abstract | Publisher Full Text

6. Abrahao R, Perdomo S, Pinto LFR, *et al.*: **Predictors of Survival After Head and Neck Squamous Cell Carcinoma in South America: The InterCHANGE Study.** *JCO Glob Oncol.* Mar 2020; **6**: 486–499.
PubMed Abstract | Publisher Full Text | Free Full Text

7. IARC/WHO: **HEADSPACE Data Center.**
Reference Source

8. Ingarfield K, McMahon AD, Hurley K, *et al.*: **Inequality in survival of people with head and neck cancer: Head and Neck 5000 cohort study.** *Head Neck.* Apr 2021; **43**(4): 1252–1270.
PubMed Abstract | Publisher Full Text

9. Ness AR, Waylen A, Hurley K, *et al.*: **Establishing a large prospective clinical cohort in people with head and neck cancer as a biomedical resource: head and neck 5000.** *BMC Cancer.* Dec 17 2014; **14**: 973.
PubMed Abstract | Publisher Full Text | Free Full Text

10. Hadji M, Rashidian H, Marzban M, *et al.*: **The Iranian Study of Opium and Cancer (IROPICAN): Rationale, Design, and Initial Findings.** *Arch. Iran. Med.* Mar 1 2021; **24**(3): 167–176.
Publisher Full Text

11. Mena M, Taberna M, Tous S, *et al.*: **Double positivity for HPV-DNA/p16(ink4a) is the biomarker with strongest diagnostic accuracy and prognostic value for human papillomavirus related oropharyngeal cancer patients.** *Oral Oncol.* Mar 2018; **78**: 137–144.
PubMed Abstract | Publisher Full Text

12. Divaris K, Olshan AF, Smith J, *et al.*: **Oral health and risk for head and neck squamous cell carcinoma: the Carolina Head and Neck Cancer Study.** *Cancer Causes Control.* Apr 2010; **21**(4): 567–575.
PubMed Abstract | Publisher Full Text | Free Full Text

13. Troy JD, Grandis JR, Youk AO, *et al.*: **Childhood passive smoke exposure is associated with adult head and neck cancer.** *Cancer Epidemiol.* Aug 2013; **37**(4): 417–423.
PubMed Abstract | Publisher Full Text | Free Full Text

14. Kachuri L, Saarela O, Bojesen SE, *et al.*: **Mendelian Randomization and mediation analysis of leukocyte telomere length and risk of lung and head and neck cancers.** *Int. J. Epidemiol.* Jun 1 2019; **48**(3): 751–766.
PubMed Abstract | Publisher Full Text | Free Full Text

15. Thomas S, Carroll JC, Brown MC, *et al.*: **Nicotine dependence as a risk factor for upper aerodigestive tract (UADT) cancers: A mediation analysis.** *PLoS One.* 2020; **15**(8): e0237723.
PubMed Abstract | Publisher Full Text | Free Full Text

16. Conlon MSC, Santi SA, Meigs ML, *et al.*: **Cigarette-smoking characteristics and interest in cessation in patients with head-and-neck cancer.** *Curr. Oncol.* Oct 2020; **27**(5): e478–e485.
PubMed Abstract | Publisher Full Text | Free Full Text

17. Santi SA, Conlon MSC, Meigs ML, *et al.*: **Rates of Smoking Cessation at 6 and 12 Months after a Clinical Tobacco Smoking Cessation Intervention in Head and Neck Cancer Patients in Northern Ontario, Canada.** *Curr. Oncol.* Mar 2 2022; **29**(3): 1544–1558.
PubMed Abstract | Publisher Full Text | Free Full Text

18. Gormley M, Adhikari A, Dudding T, *et al.*: **VOYAGER: an international consortium investigating the role of human papilloma virus and genetics in oral and oropharyngeal cancer risk and survival.** *medRxiv.* Feb 22 2025.
PubMed Abstract | Publisher Full Text | Free Full Text

19. Clements K, Thapa A, Cowell A, *et al.*: **Impact of the COVID-19 pandemic on stage and incidence of head and neck cancer: A rapid review and meta-analysis.** *Clin. Otolaryngol.* Jul 2024; **49**(4): 363–375.
PubMed Abstract | Publisher Full Text

20. Popovic M, Fiano V, Moirano G, *et al.*: **The Impact of the COVID-19 Pandemic on Head and Neck Cancer Diagnosis in the Piedmont Region, Italy: Interrupted Time-Series Analysis.** *Front. Public Health.* 2022; **10**: 809283.
PubMed Abstract | Publisher Full Text | Free Full Text

21. Clements K, Cowell A, White G, *et al.*: **The COVID-19 pandemic has not changed stage at presentation nor treatment patterns of head and neck cancer: A retrospective cohort study.** *Clin. Otolaryngol.* Jul 2023; **48**(4): 587–594.
PubMed Abstract | Publisher Full Text | Free Full Text

22. Creaney G, McMahon AD, Ross AJ, *et al.*: **Head and neck cancer in the UK: what was the stage before COVID-19? UK cancer registries analysis (2011-2018).** *Br. Dent. J.* Nov 2022; **233**(9): 787–793.
PubMed Abstract | Publisher Full Text | Free Full Text

23. Arboleda LPA, de Carvalho GB, Santos-Silva AR, *et al.*: **Squamous Cell Carcinoma of the Oral Cavity, Oropharynx, and Larynx: A Scoping Review of Treatment Guidelines Worldwide.** *Cancers (Basel).* Sep 3 2023; **15**.
Publisher Full Text

24. Matos LL, Kowalski LP, Chaves ALF, *et al.*: **Latin American Consensus on the Treatment of Head and Neck Cancer.** *JCO Glob Oncol.* Apr 2024; **10**: e2300343.
PubMed Abstract | Publisher Full Text

25. Gormley M, Creaney G, Schache A, *et al.*: **Reviewing the epidemiology of head and neck cancer: definitions, trends and risk factors.** *Br. Dent. J.* Nov 2022; **233**(9): 780–786.
PubMed Abstract | Publisher Full Text | Free Full Text

26. Smith CDL, McMahon AD, Lyall DM, *et al.*: **Development and external validation of a head and neck cancer risk prediction model.** *Head Neck.* Sep 2024; **46**(9): 2261–2273.
Publisher Full Text

27. Creaney G, de Aquino Goulart M, McMahon A, *et al.*: **Advanced Stage Head and Neck Cancer Diagnosis: HEADSpAcE Consortium Health Systems Benchmarking Survey.** *Head Neck.* Jul 2025; **47**(7): 1977–1988.
PubMed Abstract | Publisher Full Text | Free Full Text

28. Ferreiro-Iglesias A, McKay JD, Brenner N, *et al.*: **Germline determinants of humoral immune response to HPV-16 protect against oropharyngeal cancer.** *Nat. Commun.* Oct 12 2021; **12**: 5945.
PubMed Abstract | Publisher Full Text | Free Full Text

29. ACL L, Hueniken K, Pienkowski M, *et al.*: **Performance of 8 Smoking Metrics for Modeling Survival in Head and Neck Squamous Cell Carcinoma.** *JAMA Otolaryngol. Head Neck Surg.* Apr 1 2025; **151**(4): 360–370.
PubMed Abstract | Publisher Full Text | Free Full Text

30. Thakral A, Lee JJ, Hou T, *et al.*: **Smoking and alcohol by HPV status in head and neck cancer: a Mendelian randomization study.** *Nat. Commun.* Sep 7 2024; **15**: 7835.
PubMed Abstract | Publisher Full Text | Free Full Text

31. Budhathoki S, Diergaarde B, Liu G, *et al.*: **A risk prediction model for head and neck cancers incorporating lifestyle factors, HPV serology and genetic markers.** *Int. J. Cancer.* May 15 2023; **152**(10): 2069–2080.
PubMed Abstract | Publisher Full Text | Free Full Text

32. McKay JD, Truong T, Gaborieau V, *et al.*: **A genome-wide association study of upper aerodigestive tract cancers conducted within the INHANCE consortium.** *PLoS Genet.* Mar 2011; **7**(3): e1001333.
PubMed Abstract | Publisher Full Text | Free Full Text

33. Ebrahimi E, Sangphukieo A, Park HA, *et al.*: **Cross-ancestral GWAS identifies 29 variants across head and neck cancer subsites.** *Nat. Commun.* Oct 2 2025; **16**: 8787.
PubMed Abstract | Publisher Full Text | Free Full Text

34. Lydiatt WM, Patel SG, O'Sullivan B, *et al.*: **Head and Neck cancers-major changes in the American Joint Committee on cancer eighth edition cancer staging manual.** *CA Cancer J. Clin.* Mar 2017; **67**(2): 122–137.
PubMed Abstract | Publisher Full Text

35. **Obesity: preventing and managing the global epidemic. Report of a WHO consultation.** *World Health Organ. Tech. Rep. Ser.* 2000. **894**: i–xii. 1–253.
PubMed Abstract

36. Weller D, Vedsted P, Rubin G, *et al.*: **The Aarhus statement: improving design and reporting of studies on early cancer diagnosis.** *Br. J. Cancer.* Mar 27 2012; **106**(7): 1262–1267.
PubMed Abstract | Publisher Full Text | Free Full Text

37. Creaney G, Lyra Gonzalez I, Cuello M, *et al.*: **Health System Factors in Head and Neck Cancer Diagnosis: A HEADSpAcE Consortium Qualitative Study in Glasgow and Montevideo.** *Head Neck.* May 10 2026; **48**: 2849–2862.
PubMed Abstract | Publisher Full Text | Free Full Text

38. IARC: **HEADSpAcE – Head and Neck Cancer Epidemiology and Genomic Studies on Prognosis and Survival.**
Reference Source

39. Virani S: **Data from: HEADSpAcE Data Center: source data for "The HEADSpAcE Data Center: An International Data and Biospecimen Resource for Head and Neck Cancer Translational Research".** *Zenodo.* 2026.
Publisher Full Text

40. Brennan P, McKay J, Diergaarde B, *et al.*: **Data from: OncoArray: Oral and Pharynx Cancer.** 2022. dbGaP. Deposited 2022-08-22.
Reference Source

41. Viranis SBP, Olshan AF, Hung R, *et al.*; **Data from: GWA Study of Oral Cavity, Pharynx and Larynx Cancers in European, and North and South American Populations - CIDR.** 2024. dbGaP. Deposited 2024-04-15.
Reference Source

42. Torrens L, Moody S, de Carvalho AC, *et al.*: **The complexity of tobacco smoke-induced mutagenesis in head and neck cancer.** *Nat. Genet.* Apr 2025; **57**(4): 884–896.
PubMed Abstract | Publisher Full Text | Free Full Text

43. Waterboer T, Sehr P, Michael KM, *et al.*: **Multiplex human papillomavirus serology based on in situ-purified glutathione s-transferase fusion proteins.** *Clin. Chem.* Oct 2005; **51**(10): 1845–1853.
PubMed Abstract | Publisher Full Text

44. Waterboer T, Sehr P, Pawlita M: **Suppression of non-specific binding in serological Luminex assays.** *J. Immunol. Methods.* Feb 20 2006; **309**(1–2): 200–204.
Publisher Full Text

45. Fakhry C, Lacchetti C, Rooper LM, *et al.*: **Human Papillomavirus Testing in Head and Neck Carcinomas: ASCO Clinical Practice Guideline Endorsement of the College of American Pathologists Guideline.** *J. Clin. Oncol.* Nov 1 2018; **36**(31): 3152–3161.
PubMed Abstract | Publisher Full Text