Causal attribution of human papillomavirus genotypes to invasive cervical cancer worldwide: a systematic analysis of the global literature

Feixue Wei, Damien Georges, Irene Man, Iacopo Baussano, Gary M Clifford

Summary

Background Understanding the proportion of invasive cervical cancer (ICC) caused by different human papillomavirus (HPV) genotypes can inform primary (ie, vaccination) and secondary (ie, screening) prevention efforts that target specific HPV genotypes. However, using the global literature to estimate population attributable fractions (AFs) requires a methodological framework to address HPV genotype-specific causality from aggregated data. We aimed to estimate the proportion of ICC caused by different HPV genotypes at the global, regional, and national level.

Methods This systematic review identified studies reporting HPV genotype-specific prevalence in ICC or people with normal cervical cytology. We searched PubMed, Embase, Scopus, and Web of Science up to Feb 29, 2024, using the search terms "cervix" and "HPV", with no language restrictions. Odds ratios (ORs) were estimated by comparing HPV genotype-specific prevalence between HPV-positive ICC and normal cervical cytology with logistic regression models, adjusting for region, year of paper publication, and HPV primer or test. HPV genotypes with a lower bound to the 95% CI of the OR greater than 1·0 were judged as causal to ICC. Corresponding regional genotype-specific AFs were calculated as regional HPV prevalence in ICC multiplied by (1–[1/OR]) and were proportionally adjusted to total 100%. Global AFs were calculated from regional AFs weighted by number of regional ICC cases in 2022 (GLOBOCAN).

Findings The systematic review identified 1174 studies with 111 902 cases of HPV-positive ICC and 2 755 734 of normal cervical cytology. 17 HPV genotypes were considered causal to ICC, with ORs ranging widely from 48·3 (95% CI 45·7–50·9) for HPV16 to 1·4 (1·2–1·7) for HPV51. HPV16 had the highest global AF (61·7%), followed by HPV18 (15·3%), HPV45 (4·8%), HPV33 (3·8%), HPV58 (3·5%), HPV31 (2·8%), and HPV52 (2·8%). Remaining causal genotypes (HPV35, 59, 39, 56, 51, 68, 73, 26, 69, and 82) had a combined global AF of 5·3%. AFs for HPV16 and 18 and HPV16, 18, 31, 33, 45, 52, and 58 combined were lowest in Africa (71·9% and 92·1%, respectively) and highest in central, western, and southern Asia (83·2% and 95·9%, respectively). HPV35 had a higher AF in Africa (3·6%) than other regions (0·6–1·6%).

Interpretation This study provides a comprehensive global picture of HPV genotype-specific AFs in ICC, before the influence of HPV vaccination. These data can inform HPV genotype-specific vaccination and screening strategies to reduce the burden of ICC.

Funding EU Horizon 2020 Research and Innovation Programme.

Lancet 2024; 404: 435–44

See **Comment** page 406

International Agency for Research on Cancer (IARC/WHO), Early Detection, Prevention and Infections Branch, Lyon, France
(F Wei PhD, D Georges MSc, I Man PhD, I Baussano MD, G M Clifford PhD)

Correspondence to:
Dr Gary M Clifford, International Agency for Research on Cancer (IARC/WHO), Early Detection, Prevention and Infections Branch, Lyon 69007, France
cliffordg@iarc.who.int

Introduction

Invasive cervical cancer (ICC) represents a significant global health burden, with 662 301 incident cases and 348 874 deaths in 2022,[1] for which carcinogenic human papillomavirus (HPV) genotypes are considered the necessary cause.[2] In response to this high preventable burden, WHO launched its Global Strategy for cervical cancer elimination in 2022, emphasising the vital role of HPV vaccination, screening, and treatment.[3]

More than 200 HPV genotypes have been characterised, with highly variable carcinogenicity and prevalence (both in the general population and among individuals with ICC).[4,5] Therefore, accurate epidemiological assessment of causality of individual HPV genotypes in ICC and, most importantly, of their population attributable fraction (AF), is crucial for the development and impact predictions of vaccines and screening tests, which target and test for specific HPV genotypes. For ICC, the AF is the proportion of ICC cases that would not have occurred if an HPV genotype had been totally absent from the population. However, attribution of genotype-specific causality presents multiple difficulties, particularly for rarer HPV genotypes of lower carcinogenicity, as the simple detection of an HPV genotype in ICC might be due to confounding or misclassification with another HPV genotype because of common sexual transmission routes; truly causal genotypes and transient HPV co-infections require distinction.[4]

Research in context

Evidence before this study

Invasive cervical cancer (ICC) is widely acknowledged to be caused by a set of carcinogenic human papillomavirus (HPV) genotypes (13 genotypes were evaluated as carcinogenic by the last International Agency for Research on Cancer monograph working group in 2009). We conducted a comprehensive literature search in PubMed, Embase, Scopus, and Web of Science on March 5, 2024, to identify research related to human papillomavirus (HPV) population attribution fraction (AF) in ICC worldwide. The search covered studies published from database inception to Feb 29, 2024, using the search terms "cervix" and "HPV", with no language restrictions.

The largest single multicentre study reporting geographical differences in HPV genotype-specific AFs was published in 2009 and included 10 575 ICC cases worldwide. However, the representation of ICC cases in high-burden areas, such as Africa (n=544) and some countries in Asia, namely India (n=252) and China (n=200), was low, and no individual national-level data—nor global estimates weighted for regional ICC burden—were presented. Although several systematic reviews and pooled analyses using aggregated data from scientific literature on ICC up to 2010 have also provided valuable insights into HPV genotype-specific prevalence in ICC on both global and regional scales, they did not attempt to estimate HPV genotype-specific AFs. Addressing this gap requires a methodology that can overcome the various challenges in distinguishing causal HPV genotypes from transient HPV co-infections that can be applied to aggregated data from the scientific literature.

Added value of this study

This study collated the entire current and relevant scientific literature—which has expanded considerably in the last decade—to provide a comprehensive global and regional overview of HPV genotype-specific AFs in ICC. The results include aggregated HPV prevalence data from over 111 000 ICC cases in 121 countries. The novel methodology was based on odds ratios (ORs) comparing HPV genotype-specific prevalence between people with HPV-positive ICC and people with normal cervical cytology to evaluate which HPV genotypes are causal and, if causal, what fraction of ICC is attributable to individual HPV genotypes. This approach judged 17 HPV genotypes as causal, but with wide variation in carcinogenicity, from HPV16 representing by far the highest risk (OR 48·3) down to HPV51 showing the lowest risk (1·4). By applying these risks to calculate AFs, HPV16 and 18 consistently accounted for approximately three-quarters of ICC cases across all global regions, with HPV31, 33, 45, 52, and 58 contributing towards an additional 15–20%. The remaining ten attributable genotypes (HPV35, 59, 39, 56, 51, 68, 73, 26, 69, and 82) were responsible for approximately 5% of ICC cases worldwide, with some notable small regional variations, particularly a higher AF for HPV35 in Africa.

Implications of all the available evidence

These data can inform ICC prevention strategies that target specific HPV genotypes to accelerate cervical cancer elimination, both through primary prevention (HPV vaccination), and secondary prevention (HPV-based screening). The study highlights the significance of existing vaccines in global ICC prevention while also highlighting some regional and national specificities, such as increased HPV35 importance to ICC in Africa, which should be included in future HPV vaccines to reduce disparities with other regions, and a high HPV16 and 18 AF in India. These data therefore allow us to predict with unprecedented accuracy the expected impact of HPV vaccination, using current and future HPV vaccines, on ICC burden at global, regional, and national levels.

Moreover, this research provides a valuable new perspective on the implications of testing for individual, or groups of, HPV genotypes for clinical decision making in cervical cancer screening programmes. The eight HPV genotypes that had the highest AFs (HPV16, 18, 31, 33, 35, 45, 52, and 58) were also the most carcinogenic and are therefore clear priority targets for inclusion in HPV-based screening tests. Other HPV genotypes, although each individually responsible for a small fraction of ICC, showed a sliding scale of lower positive predictive value and their inclusion in cervical cancer screening tests incrementally reduces screening efficiency and cost-effectiveness.

With respect to causality, in 2009, an International Agency for Research on Cancer (IARC) Working Group classified HPV16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, and 59 as carcinogenic (Group 1) to humans, and HPV68 as probably carcinogenic (Group 2A), according to an algorithm that compared HPV genotype-specific prevalence between ICC cases and people with normal cervical cytology (as a proxy for the background HPV prevalence in female individuals), using aggregated data.[4,6] These 13 Group 1/2A genotypes cluster together in an evolutionary clade that includes α-papillomavirus species groups 5, 6, 7, and 9. Other genotypes in this high-risk clade were classified as possibly carcinogenic (Group 2B), based solely on phylogenetic relatedness.[4,6]

Over the past decade, there has been an increase in worldwide reports on HPV genotype-specific prevalence in ICC and, following widespread implementation of HPV-based cervical cancer screening programmes, an even greater increase in publications on HPV genotype-specific prevalence among people with normal cervical cytology. Leveraging this information, we conducted a systematic review to judge the carcinogenicity of individual HPV genotypes in the high-risk clade, and to provide a comprehensive summary of the AF of individual HPV genotypes in ICC at global, regional, and national levels.

Methods

Search strategy and selection criteria

In this systematic analysis, we updated a previous systematic review on cervical HPV infection across the spectrum of cervical disease,[7,8] by searching for publications in PubMed that were published between Dec 1, 2011, and Feb 29, 2024. The search was also extended to Embase, Scopus, and Web of Science, with publication dates from database inception to Feb 29, 2024, and no language restrictions. The search terms "cervix" and "HPV" were used across databases to identify relevant literature (appendix pp 2–5). References from key studies and reviews were also scanned.

We included studies reporting genotype-specific HPV infection, detected by broad-spectrum PCR-based tests from cervical exfoliated cells or biopsies, in people with normal cervical cytology or ICC (including squamous cell carcinoma, adeno or adenosquamous carcinoma, and cervical cancer of unspecified genotype, excluding carcinoma in situ). If more than one Article presented results from the same study, we selected the Article with the most comprehensive information. Two authors (FW and GMC) independently selected studies for inclusion and extracted aggregated data from published reports. Discrepancies were resolved by discussion between these authors. Aggregated data from key studies were requested from the authors if such information was not available in the original publications.

Data analysis

We extracted data on first author; year of paper publication; journal name; country of recruitment; HPV DNA source (exfoliated cells or biopsies); HPV DNA PCR primer or test—[PG]MY09/11/Linear Array, GP5(+)/6(+), SPF10/LIPA, COBAS, GenoArray, or others; genotype-specific and overall HPV DNA prevalence according to cervical cytopathological diagnosis (normal cervical cytology or ICC); and HIV status. Notably, not all studies tested for all HPV genotypes, so denominators of prevalence vary by HPV genotype and, because extracted data are aggregated, crude HPV prevalence represents that in single and multiple infections.

The primary outcome of this study was the AFs for individual genotypes in ICC. This measure was calculated from secondary outcomes of pooled HPV genotype-specific prevalence in people with ICC and in people with normal cervical cytology.

Pooled HPV genotype-specific prevalence and corresponding 95% CIs were estimated by using logistic regression models at national, regional, and global levels. Odds ratios (ORs) with 95% CIs were calculated by comparing HPV genotype-specific prevalence in HPV-positive ICC cases versus that in people with normal cervical cytology,[4] for all individual HPV genotypes in the high-risk clade (ie, IARC Group 1/2A/2B),[6] using random-effect logistic regression models, adjusted for HPV primer or test, year of paper publication, and region. Regions were stratified, according to UN subregions, as Africa; eastern and southeastern Asia; central, western, and southern Asia; Europe; North America; Latin America and the Caribbean; and Oceania. Main analyses compared genotype-specific HPV prevalence in all cancer cases (from either exfoliated cells or biopsies), with people with normal cervical cytology. A sensitivity analysis was performed, restricted only to cancer cases where HPV prevalence was assessed in exfoliated cells only.

See Online for appendix

	People with invasive cervical cancer (n=111 902)	People with normal cervical cytology (n=2 755 734)
Geographical region		
Eastern and southeastern Asia	54 768 (48·9%)	1 478 814 (53·7%)
Central, western, and southern Asia	7746 (6·9%)	38 727 (1·4%)
Europe	22 743 (20·3%)	856 630 (31·1%)
Latin America and the Caribbean	11 460 (10·2%)	115 230 (4·2%)
North America	4333 (3·9%)	232 964 (8·5%)
Africa	8834 (7·9%)	27 668 (1·0%)
Oceania	2018 (1·8%)	5701 (0·2%)
Year of paper publication		
1986–2000	7930 (7·1%)	19 204 (0·7%)
2001–10	33 829 (30·2%)	279 646 (10·2%)
2011–20	54 569 (48·8%)	1 779 094 (64·6%)
2021–24*	15 574 (13·9%)	677 790 (24·6%)
HIV status		
Negative or unknown	110 283 (98·6%)	2 745 023 (99·6%)
Positive†	1619 (1·5%)	10 711 (0·4%)
HPV DNA source		
Cells	46 329 (41·4%)	2 755 734 (100·0%)
Biopsies	65 573 (58·6%)	0
HPV primer or test		
(PG)MY09/11/Linear Array	42 207 (37·7%)	866 154 (31·4%)
SPF10/LIPA	24 144 (21·6%)	109 656 (4·0%)
GP5(+)/6(+)	14 204 (12·7%)	196 533 (7·1%)
GenoArray	5332 (4·8%)	177 937 (6·5%)
COBAS	1118 (1·0%)	625 733 (22·7%)
Others	24 897 (22·3%)	779 721 (28·3%)
Histological type		
Squamous cell carcinoma	44 773 (40·0%)	..
Adeno or adenosquamous carcinoma	9166 (8·2%)	..
Unspecified type	57 963 (51·8%)	..

Data are n (%). HPV=human papillomavirus. *Until February, 2024. †Of whom, 1154 (71·2%) people with invasive cervical cancer and 2935 (27·4%) people with normal cervical cytology were from Africa.

Table 1: **Baseline characteristics**

	HPV species	IARC 100B group	People with invasive cervical cancer		People with normal cervical cytology		OR (95% CI)*
			n	Prevalence (%; 95% CI)	n	Prevalence (%; 95%CI)	
HPV16	α9	Group 1	111 812	63·9% (63·6–64·2)	2 509 575	2·1% (2·1–2·2)	48·3 (45·7–50·9)
HPV18	α7	Group 1	106 425	16·8% (16·6–17·1)	2 441 188	0·9% (0·9–0·9)	12·5 (11·6–13·6)
HPV45	α7	Group 1	87 789	4·7% (4·6–4·9)	1 745 145	0·5% (0·5–0·5)	6·6 (5·8–7·4)
HPV33	α9	Group 1	93 978	4·9% (4·7–5·0)	1 678 191	0·7% (0·7–0·7)	6·4 (5·8–7·1)
HPV31	α9	Group 1	93 574	4·0% (3·8–4·1)	1 955 611	0·9% (0·9–0·9)	3·9 (3·5–4·3)
HPV58	α9	Group 1	88 340	6·0% (5·8–6·1)	1 846 314	1·4% (1·4–1·4)	3·9 (3·5–4·2)
HPV52	α9	Group 1	89 409	4·9% (4·8–5·0)	1 755 085	2·3% (2·2–2·3)	3·1 (2·8–3·4)
HPV59	α7	Group 1	80 619	1·6% (1·6–1·7)	1 595 415	0·6% (0·6–0·6)	2·6 (2·2–3·0)
HPV26	α5	Group 2B	40 507	0·3% (0·3–0·4)	244 096	<0·1% (<0·1–<0·1)	2·6 (1·5–4·0)
HPV69	α5	Group 2B	24 429	0·3% (0·2–0·3)	196 666	0·1% (0·1–0·1)	2·4 (1·4–3·7)
HPV35	α9	Group 1	79 991	2·0% (1·9–2·1)	1 641 402	0·5% (0·4–0·5)	2·3 (2·0–2·7)
HPV39	α7	Group 1	79 120	1·6% (1·5–1·7)	1 643 824	0·9% (0·9–1·0)	1·8 (1·6–2·1)
HPV73	α11	Group 2B	49 642	0·7% (0·6–0·7)	784 992	0·2% (0·2–0·2)	1·8 (1·3–2·4)
HPV68	α7	Group 2A	71 442	1·2% (1·1–1·3)	1 564 316	0·9% (0·9–0·9)	1·7 (1·4–2·0)
HPV56	α6	Group 1	78 733	1·5% (1·4–1·6)	1 603 864	0·8% (0·8–0·8)	1·6 (1·3–1·8)
HPV82	α5	Group 2B	47 638	0·3% (0·3–0·4)	852 293	0·2% (0·2–0·2)	1·5 (1·0–2·0)
HPV51	α5	Group 1	76 809	1·5% (1·4–1·6)	1 664 411	1·1% (1·1–1·1)	1·4 (1·2–1·7)
HPV34	α11	Group 2B	23 271	0·1% (0·1–0·2)	132 821	<0·1% (<0·1–<0·1)	1·4 (0·6–3·0)
HPV30	α6	Group 2B	13 608	0·3% (0·2–0·4)	62 505	0·2% (0·2–0·2)	1·4 (0·4–3·1)
HPV53	α6	Group 2B	64 012	1·4% (1·3–1·5)	1 455 060	1·2% (1·2–1·2)	1·1 (0·9–1·3)
HPV67	α9	Group 2B	29 527	0·3% (0·3–0·4)	167 329	0·3% (0·3–0·3)	1·0 (0·6–1·6)
HPV66	α6	Group 2B	69 838	0·9% (0·8–1·0)	1 452 369	0·7% (0·7–0·7)	0·9 (0·7–1·1)
HPV70	α7	Group 2B	42 335	0·4% (0·4–0·5)	251 855	0·8% (0·8–0·8)	0·5 (0·4–0·7)
HPV85	α7	Group 2B	3663	0·0% (0·0–0·2)	54 489	0·2% (0·2–0·2)	0·2 (0·0–0·9)

HPV=human papillomavirus. ICC=invasive cervical cancer. IARC=International Agency for Research on Cancer. OR=odds ratio. *ORs of HPV prevalence in people with HPV-positive invasive cervical cancer and people with normal cervical cytology were separately estimated for each HPV type in IARC Monograph 100B Group 1, 2A, and 2B (except for HPV97 [of 1433 individuals with invasive cervical cancer, one individual was positive and of 4491 people with normal cervical cytology, one person was positive] due to the low prevalence that prevented the models from converging), by use of a logistic regression model with adjustment for region, HPV primer or test, and year of paper publication.

Table 2: **HPV genotype-specific prevalence in people with ICC and people with normal cervical cytology**

We defined causal genotypes as those with a lower bound for the 95% CI of the OR greater than 1·0. For these genotypes, the corresponding regional AFs were calculated by multiplying regional HPV genotype-specific prevalence in HPV-positive ICC cases by (1−[1/OR]),[9] and proportionally adjusting the resulting genotype-specific AFs to total 100% (while not allowing any AF to exceed the crude HPV prevalence in HPV-positive ICC for that genotype). We then estimated global AFs by weighting regional-level AFs according to number of regional ICC cases in GLOBOCAN for the year 2022.[1] For countries contributing more than 1000 HPV-positive ICC cases, we also estimated national AFs using the same approach. For regional and national AFs, 95% credible intervals (CrIs) were derived from 2·5% and 97·5% sample quantiles of a Bayesian model, which was fitted using RStan.[10] Based upon a Markov chain Monte Carlo algorithm with the No-U-Turn sampler, this model produced AF estimates with 95% CrIs according to HPV prevalence in ICC and ORs, and accounting for the non-independence of these two measures in the calculation of the AF by considering their potential joint probability distribution. Model convergence was assessed visually and with the R-hat statistic. We also conducted sub-analyses of HPV genotype-specific AFs according to histological genotype (squamous cell carcinoma and adeno or adenosquamous carcinoma).

To validate the methodology, we analysed individual-level data on HPV infection in 4089 HPV-positive ICC cases collected by IARC in 31 countries.[11–16] We compared AF estimations from our methodology based on aggregated data with that from three alternative approaches using individual-level data on multiple HPV genotypes.[17]

All data analysis was performed using R software (version 4.1.0). This study was registered in PROSPERO, CRD42023392489.

Role of the funding source
The funders of this study had no involvement in study design, data collection, data analysis, data interpretation, or writing of the report.

Results

This systematic review identified a total of 1174 eligible studies, including 483 from a previous systematic review[7] and 691 newly identified studies (appendix p 11). After excluding 14 642 people with HPV-negative ICC, final analyses used data from 111 902 people with HPV-positive ICC and 2 755 734 people with normal cervical cytology from 121 countries (appendix p 12). The largest sample contributions came from studies in eastern and southeastern Asia, followed by Europe (table 1). Nearly half of people with ICC and 64·6% of people with normal cervical cytology were from studies published between Jan 1, 2011, and Dec 31, 2020. The vast majority (>98%) of individuals were not known to have HIV. Among people with ICC, 58·6% (65 573 of 111 902) had HPV DNA testing from biopsies and 41·4% (46 329 of 111 902) from cervical exfoliated cells. The most common HPV primer or test was (PG)MY09/11/Linear Array in both people with ICC (37·7% [42 207 of 111 902]) and people with normal cervical cytology (31·4% [866 154 of 2 755 734]). Nearly half of the people with ICC had information on histological genotype: 40·0% (44 773 of 111 902) were squamous cell carcinoma and 8·2% (9166 of 111 902) adeno or adenosquamous carcinoma.

17 HPV genotypes were identified as significant (lower bound to the 95% CI of ORs greater than 1·0) and were considered causal (table 2). All Group 1 and 2A genotypes had significant ORs, but with strong variation from 48·3 (95% CI 45·7–50·9) for HPV16 to 1·4 (1·2–1·7) for HPV51. In Group 2B, HPV26 had the highest OR (2·6, 1·5–4·0), followed by HPV69, 73, and 82 which also had significant ORs. Other genotypes were not considered causal.

HPV16 had the highest global AF (61·7%), followed by HPV18 (15·3%), HPV45 (4·8%), HPV33 (3·8%), HPV58 (3·5%), HPV31 (2·8%), and HPV52 (2·8%; figure 1). The remaining ten causal genotypes (HPV35, 59, 39, 56, 51, 68, 73, 26, 69, and 82) had a combined global AF of 5·3%. For each of the 17 causal HPV genotypes, relationships between global AFs and prevalence in people with normal cervical cytology are shown in figure 2.

HPV16 and 18 were the predominant attributable genotypes in all regions, with combined AFs ranging from 71·9% in Africa to 83·2% in central, western, and southern Asia (figure 3). HPV45 was the third attributable genotype in all regions, except for eastern and southeastern Asia (HPV58; appendix pp 6–8). HPV33, 58, 31, 52, and 35 tended to be the fourth to eighth most attributable genotypes in all regions. In Africa, HPV35 had a higher AF (3·6%) than other regions (0·6–1·6%). Combined HPV16, 18, 31, 33, 45, 52, and 58 AFs ranged from 92·1% in Africa to 95·9% in central, western, and

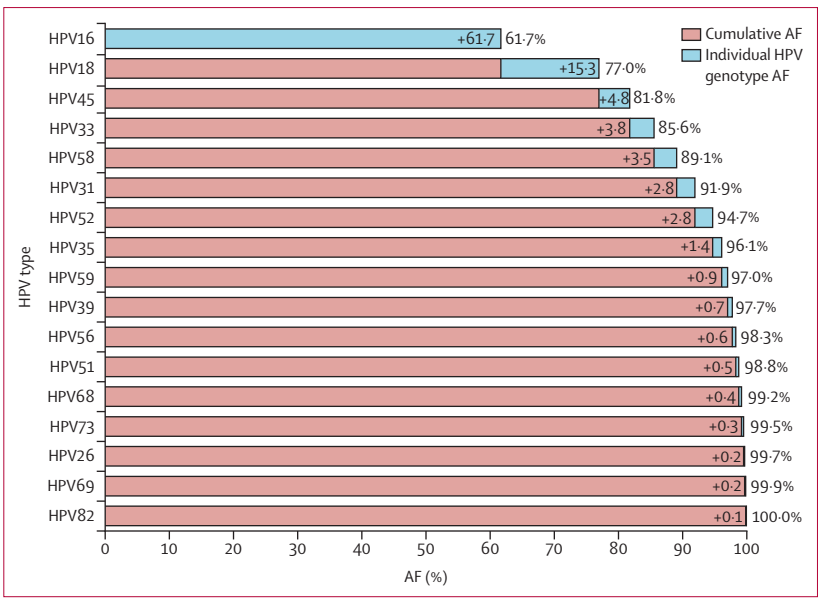

Figure 1: **Individual and cumulative HPV genotype-specific AF in invasive cervical cancer at the global level**
The number outside each bar shows the cumulative AF. AF=population attributable fraction. HPV=human papillomavirus.

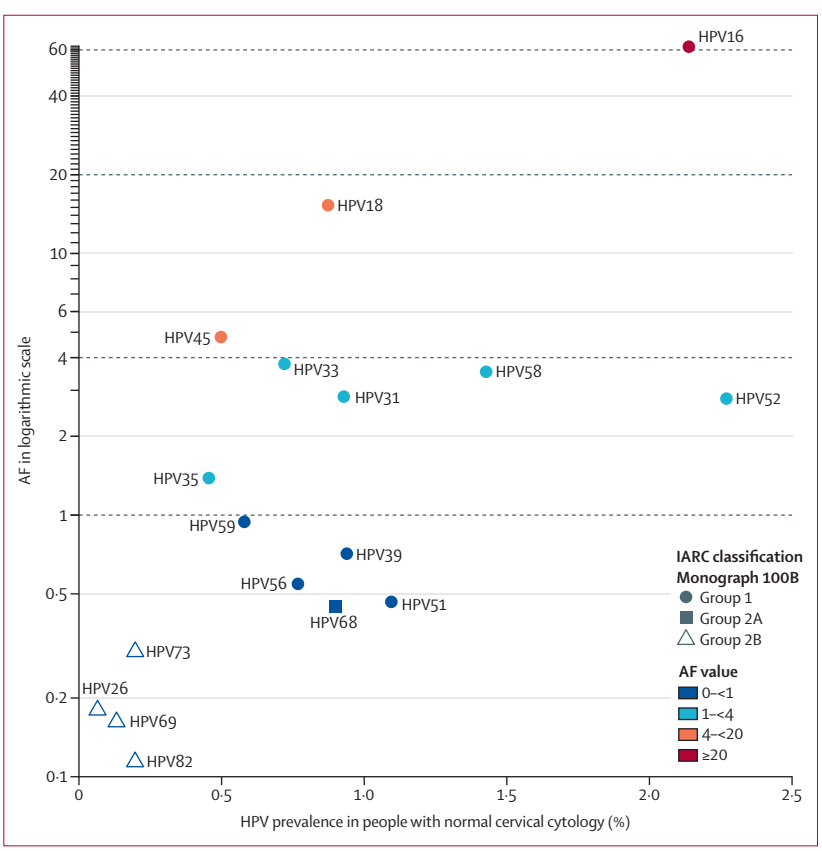

Figure 2: **HPV genotype-specific prevalence in people with normal cervical cytology *vs* corresponding AF in invasive cervical cancer**
The y-axis is in logarithmic scale. Dashed gridlines at the cut-offs of 1, 4, 20, and 60 represent grouping of HPV genotypes by AF. AF=population attributable fraction. HPV=human papillomavirus. IARC=International Agency for Research on Cancer.

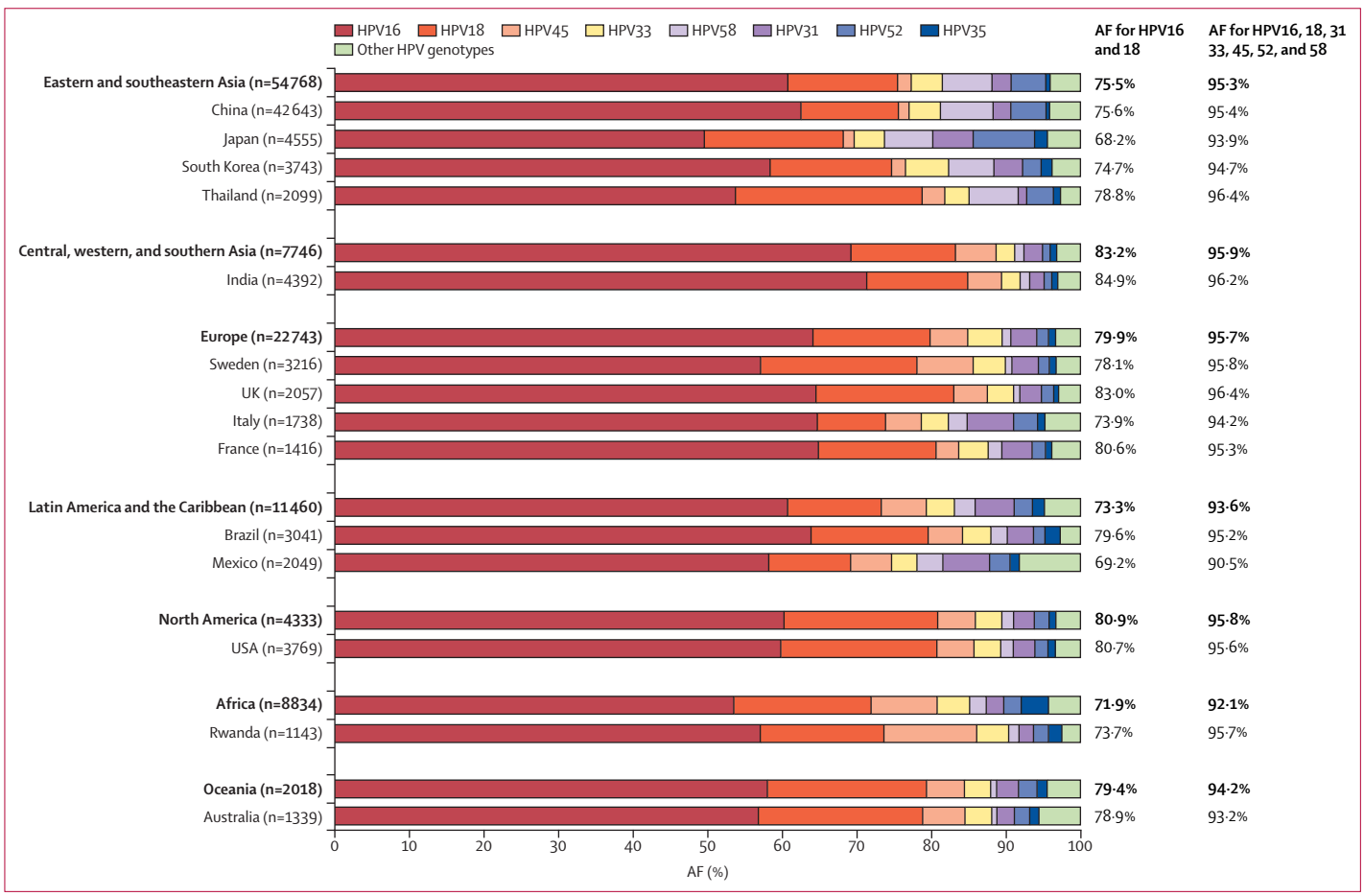

Figure 3: **HPV individual genotype-specific AF at regional and selected national levels**
Individual countries listed had more than 1000 individuals with HPV-positive invasive cervical cancer. The other HPV genotypes category represents the sum of AFs for HPV59, 39, 56, 51, 68, 73, 26, 69, and 82. AF=population attributable fraction. HPV=human papillomavirus.

southern Asia. In each of the countries with more than 1000 people with HPV-positive ICC, HPV16 and HPV18 predominated, and the combined AF of HPV16, 18, 31, 33, 45, 52, and 58 exceeded 90% (figure 3).

HPV16 contributed a higher AF in squamous cell carcinoma compared with adeno or adenosquamous carcinoma globally (63·7% *vs* 46·4%), a pattern consistent across all regions. Other attributable genotypes within the same species as HPV16 (ie, HPV31, 33, 35, 52, and 58) also had higher AFs in squamous cell carcinoma than adeno or adenosquamous carcinoma. Conversely, HPV18 had a higher global AF in adeno or adenosquamous carcinoma compared with squamous cell carcinoma (38·5% *vs* 13·2%). HPV16 and HPV18 remained the most attributable genotypes in both squamous cell carcinoma and adeno or adenosquamous carcinoma, accounting for 77·0% of squamous cell carcinoma cases and 84·9% of adeno or adenosquamous carcinoma cases worldwide, with higher AFs observed for adeno or adenosquamous carcinoma in all regions (figure 4). The combined AF for HPV16, 18, 31, 33, 45, 52, and 58 was similar in squamous cell carcinoma (94·9%) and adeno or adenosquamous carcinoma (96·9%) worldwide, with little regional variation.

In a validation exercise of 4089 cases of HPV-positive ICC with individual-level information, the study approach for calculating HPV genotype-specific AFs using HPV prevalence from aggregated data yielded similar results to using alternative approaches making use of full HPV typing information at an individual level (appendix p 9). The combined AF for HPV16 and HPV18 was 77·0% using aggregated data, and the relative estimation was 75·7% using a single infection approach, 75·9% using the hierarchical approach, and 75·5% using the proportional approach making use of full HPV genotyping information at an individual level.

In a sensitivity analysis restricted to ICC with HPV DNA tested solely from cells (excluding approximately two-thirds of ICC tested from biopsies), ORs for HPV genotypes judged to be causal in the main analysis remained significant (appendix p 10), with HPV53 (OR 1·6, 95% CI 1·3–1·9) and 66 (1·4, 1·1–1·8) also showing significant, albeit low, ORs.

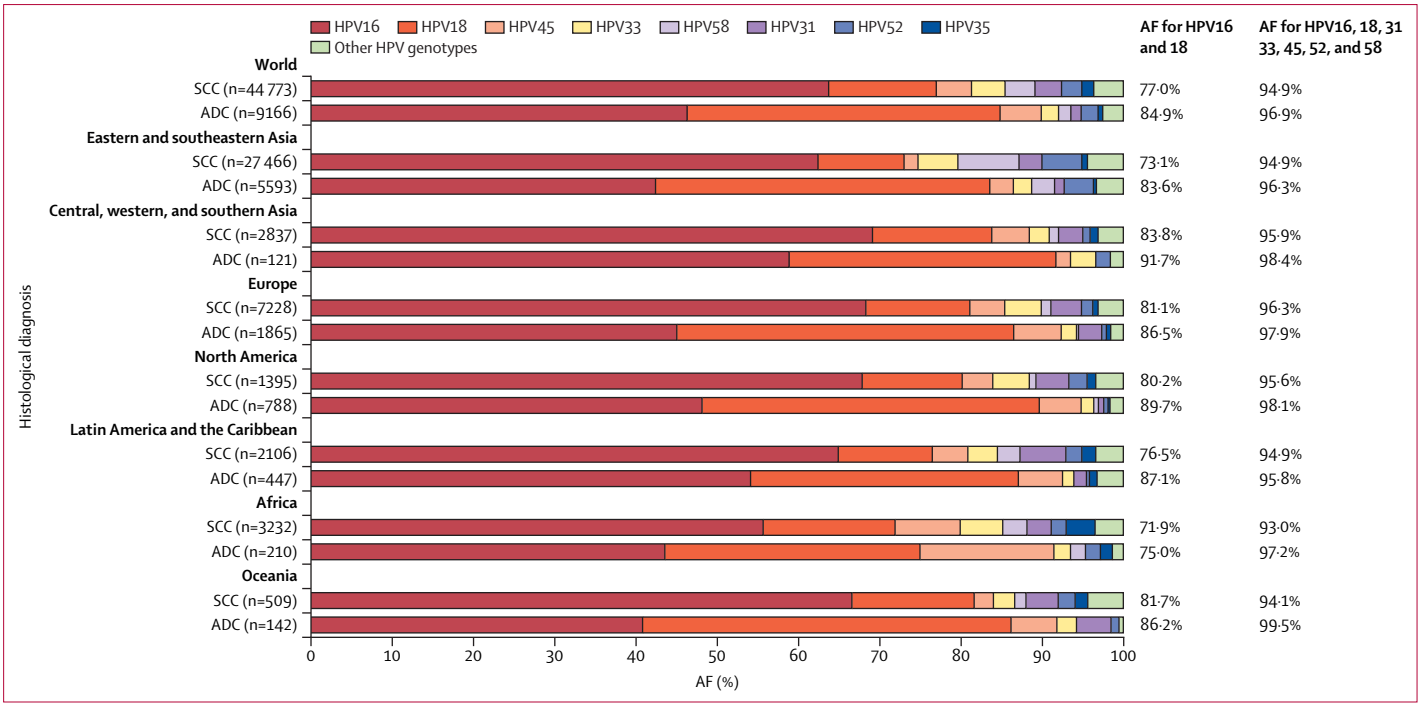

Figure 4: **HPV individual genotype-specific AF in invasive cervical cancer by histological diagnosis, at global and regional levels**
The other HPV genotypes category represents the sum of AFs for HPV59, 39, 56, 51, 68, 73, 26, 69, and 82. ADC=adeno or adenosquamous carcinoma. AF=population attributable fraction. HPV=human papillomavirus. SCC=squamous cell carcinoma.

Discussion

In this study, we present the first comprehensive global and regional picture of the proportion of ICC attributable to causal HPV genotypes. This was based on systematically aggregated data from more than 111 000 HPV-positive ICC cases in the global scientific literature. HPV16 and 18 were confirmed to consistently cause approximately three-quarters of ICC in all regions, and HPV31, 33, 45, 52, and 58 to cause an additional 15–20%. Around 5% of worldwide ICC cases were attributable to 10 remaining causal types, with some notable small regional variations, particularly for HPV35 in Africa.

Estimating genotype-specific AFs is methodologically challenging due to confounding related to common sexual transmission routes, and the difficulty in distinguishing causal HPV genotypes from transient HPV co-infections in ICC cases, particularly for rarer HPV genotypes and when working with aggregated data. Indeed, our data showed the sum of crude HPV prevalence in ICC to be more than 100% (appendix pp 6–8), due to the presence of multiple infections. We overcame this problem by adapting the methodological framework to reflect the strength of causality for individual HPV genotypes by comparing HPV prevalence in ICC cases versus that expected in people with normal cervical cytology (ORs).[4,6]

This causality judgement supported the carcinogenic potential of 17 HPV genotypes, including four Group 2B genotypes (HPV73, 26, 69, and 82) in addition to the 13 genotypes previously classified by IARC as Group 1/2A.[6] This epidemiological assessment is supported by evidence of rare biological activity of these genotypes, including single viral oncogene transcription with $p16^{INK4a}$ overexpression, pRb downregulation, and variable p53 expression, when found as a single infection among ICC cases.[18] However, although these four Group 2B genotypes were considered causal, each was only very weakly carcinogenic (low ORs) and was rare, and so less than 1% of all ICC cases were collectively attributable to these genotypes.

The approach also ensured that AFs totalled 100%, considering HPV as a necessary cause of ICC. We excluded cases of HPV-negative ICC a priori, and after estimating AFs for each HPV genotype separately, proportionally adjusted the resulting 17 individual HPV genotype-specific AFs to 100%. This approach represents a pragmatic solution to using the entirety of aggregated data in the global literature to estimate AFs at global, regional, and national levels, rather than focusing only on a low number of gold-standard studies of ICC cases. Indeed, arbitrary decisions on attribution in multiple HPV infections still need to be made even in these gold-standard ICC case studies.[19,20] Although laser capture micro-dissection is the ideal approach to truly identify the causal genotype in cases of ICC with multiple HPV infections,[21] followed by in situ hybridisation or detection of E6/E7 RNA,[18,22] these methods are time-consuming

and resource-intensive, and little representative data exist at a global scale, especially for low-income and middle-income countries, where 70% of global ICC cases occur.

HPV16 and 18 were consistently identified as the genotypes most attributable to ICC in all regions, confirming previous pooled analyses and original studies, and accounted for approximately three-quarters of global ICC. Our estimations of combined HPV16 and 18 AFs (77·0%) were slightly higher than the estimates from three (ie, single infection only; hierarchical; and proportional) approaches using individual-level data (75·5–75·9%) when applied to the same IARC data for 4089 HPV-positive ICC cases, but remained comparable (appendix p 9). Our HPV16 and 18 AFs are also somewhat higher than that reported in the largest single multicentre worldwide study[20] of 8977 individuals with HPV-positive ICC (70·8%).[20] This difference might partly be explained by our restriction of attribution to 17 causal genotypes, ignoring possible attribution to other HPV genotypes, even if present. There have been steady increases in HPV16 detection in ICC studies over time (from 60·3% in studies published between 1990 and 1999[8] and reaching 65·3% in studies published after 2020), likely due to improvements in testing protocols and overall quality of biological specimens and diagnosis (hence our OR adjustment for HPV primer or test, and year of paper publication). Finally, slightly higher global HPV16 and 18 AFs in this study might also reflect differences in geographical source of ICC cases. Regional differences can already be observed in crude HPV16 and 18 prevalence in ICC (appendix pp 6–8), and in previous pooled analyses that did not methodologically address AFs.[8]

The next most attributable genotypes were also largely consistent across all regions, and were a combination of genotypes HPV45, 33, 58, 31, and 52, with only minor exceptions.[7,11,20] HPV16, 18, 31, 33, 45, 52, and 58 collectively accounted for 94·7% of global ICC, but this estimate was again lower in Africa (92·1%). This difference could be largely attributed to HPV35, which has a substantially higher AF in Africa (3·6%) than other regions (0·6–1·6%). In the USA, HPV35 prevalence has also been reported to be higher among Black people diagnosed with ICC.[23] Inclusion of HPV35 in future vaccines would reduce these inequalities.

Our approach judged the remaining fraction of ICC attributable to non-HPV16, 18, 31, 33, 45, 52, and 58 genotypes to be 5·3%, which is somewhat lower than the 10% estimated from the largest single study to date.[20] Again, this difference might partly be explained by our attribution restriction to the remaining ten genotypes judged to be causal, thereby ignoring: (1) possible attribution of ICC to other HPV genotypes; (2) increasing HPV16 positivity over time; and (3) the regional representation of ICC cases.

Our findings offer particularly valuable insights and guidance in planning effective HPV vaccination strategies in specific countries where sufficient numbers of ICC have been genotyped for HPV. We were able to provide AF estimates for India and China, which respectively account for 19% and 23% of global cervical cancer incidence,[1] and have yet to implement national vaccination programmes. Combined HPV16 and 18 AF was higher in India (84·9%) compared with China (75·6%), but the AFs for HPV16, 18, 31, 33, 45, 52, and 58 were similar (96·2% in India and 95·4% in China). Notably, the combined AF for HPV16 and 18 was only 68·2% in Japan due to the lower crude HPV16 prevalence,[24] but the AF for combined HPV16, 18, 31, 33, 45, 52, and 58 still exceeds 90% (with a higher HPV52 AF than other analysed countries). These data are key for predicting the potential impact of current and future HPV vaccines on the ICC burden in these countries.[25] For countries not specifically presented in the current analysis, regional genotype-specific AFs are expected to offer relevant estimates.

In addition to predicting HPV vaccine impact, these data can also inform the impact of testing for individual HPV genotypes during cervical cancer screening. Indeed, although ORs presented here do not represent absolute risks for developing ICC from HPV infection, they are useful surrogates for ranking genotypes by positive predictive value for cancer. Similarly, genotype-specific AFs can be used as a surrogate for potential gains in sensitivity, and genotype-specific prevalence in people with normal cervical cytology as a surrogate for potential losses in specificity, by including a given HPV genotype into a screening test (figure 2).[26] Notably, the eight HPV genotypes with highest individual AFs (16, 18, 31, 33, 35, 45, 52, and 58) also showed the strongest positive predictive value (ie, had the highest ORs, with the slight exception of HPV35), and are therefore clear priorities for maximising screening efficiency. The remaining nine HPV genotypes showed a sliding scale of lower positive predictive value and lower AFs, and testing for them could provoke more unnecessary follow-up visits and reduce impact on cervical cancer prevention, thereby reducing screening efficiency and cost-effectiveness. This group of nine HPV genotypes include a mix of classifications as IARC Group 1, 2A, and 2B,[6] and can also be the common cause of low-grade and high-grade cervical lesions,[7] causing many unnecessary biopsies and treatments. In a population-level study in Sweden, the ranking of HPV genotypes by the number of women needed to screen and number of women needing follow-up to detect or prevent one cervical cancer,[27] as a measure of screening efficiency, offered a similar ranking to the ORs presented here. Our estimates are also complementary to genotype-specific estimates of sensitivity-to-specificity ratios for cervical intraepithelial neoplasia grade 3 (known as CIN3),[28] recognising that not all CIN3 will progress to cancer, and that CIN3 progression risk might differ by genotype.[7]

Current US Food and Drug Administration-approved HPV screening tests, such as Cervista, APTIMA, Cobas,

and HC2, include the 13 genotypes classified as Group 1 or 2A, and some include Group 2B HPV66, which was judged as non-attributable in our main analysis. Notably, HPV51 in Group 1 was common among people with normal cervical cytology but attributable to only a very small percentage of cancers, so including HPV51 in tests adversely affected specificity; HPV51 is also the cause of an important fraction of low-grade and high-grade cervical lesions.[7] Group 2B genotypes 26, 69, and 73 had higher ORs compared with several genotypes in Group 1 and 2A, but were rarer than HPV51 in people with normal cervical cytology, reducing concerns about over-referral for many people at low risk of ICC. Acceptable trade-offs in sensitivity versus specificity will depend on local strategies and resources for following up screen positives. In low-income and middle-income countries in particular, specificity and screening efficiency are important for programmatic feasibility and success. Our data also inform hierarchical grouping of HPV genotypes in tests for clinical decision making.[29]

Cross-sectional pooled analyses of HPV genotype distribution have generic limitations, including variations in age ranges and quality of diagnosis across different studies.[7] In this study, ICC cases were under-represented in certain regions, including those where the ICC burden is high (eg, Africa). We therefore adjusted for region when estimating ORs, presented AFs at the regional level, and weighted global AFs by the number of regional cases based on GLOBOCAN data. Although we presented data stratified by histological genotype where possible, highlighting well established differences in genotype-specific AFs between squamous cell carcinoma and adeno or adenosquamous carcinoma,[8,20] most ICC cases did not allow this stratification. In addition, we did not have a satisfactory method to allow weighting of regional and global AFs by histological genotype, so we pooled all cases, irrespective of histology. We also pooled HPV infection data from biopsies and cells for ICC, while acknowledging that HPV genotype-specific prevalence might differ in ICC depending on the source of HPV DNA tested,[30] and that HPV prevalence in normal cervical cytology was assessed entirely from cells. However, the vast majority of ICC were tested from biopsies, which are expected to give more specific representation of causal genotypes than cells, and it is HPV prevalence in ICC that drives AFs. AFs are ideally estimated using risk ratios, not ORs. However, there are no estimates of risk ratios for ICC by HPV genotypes from cohort studies (mainly because cervical precancers are screened for, and treated), and the use of ORs as a valid surrogate measure has been widely applied previously for cervical cancer[6] and other infectious causes of cancer.[31]

There were some HPV genotypes at the limits of causality judgements. For example, HPV30, 34, 53, and 67 had ORs greater than 1, but a lower bound of the 95% CI that was less than 1. Furthermore, in the sensitivity analysis restricted to ICC with HPV testing from cells only, HPV66 and 53 also had significant ORs. However, even if all these borderline genotypes were to be considered causal, they were only very weakly carcinogenic (ie, low ORs), and so no more than 0·2% of all ICC cases were collectively attributable to these genotypes. Yet some of these genotypes are very common in people with normal cervical cytology (collective prevalence of over 2%) and are not recommended to be included in HPV screening tests due to reductions in specificity, and unnecessary harms due to management (ie, follow-up visits, biopsies, and treatments) of people at extremely low cancer risk.

Although HIV infection is known to affect HPV-genotype distribution in ICC, based on data derived almost exclusively from Africa,[32] we were not able to take a stratified approach to calculating ORs by HIV status, due to a low sample size (only 0·4% of people with normal cervical cytology were known to have HIV). Instead, given that the proportion of ICC cases from people with HIV from Africa in the current pooled analysis (13%) was considered an under-representation of the actual fraction for this subpopulation (20%),[33] we pooled ICC cases, and people with normal cervical cytology, irrespective of HIV status. Finally, although we excluded studies specifically recruiting people who had been vaccinated for HPV, most studies did not provide information on HPV vaccination status. However, based on age at enrolment, country, and year of paper publication, HPV prevalence would not yet have been expected to be influenced by HPV vaccination.

In conclusion, by developing a pragmatic approach to collate the entirety of the relevant scientific literature, we have achieved a comprehensive global overview for the AF of ICC caused by 17 HPV genotypes. This study confirms the importance of existing vaccines for global ICC prevention, while highlighting some regional specificities, most notably HPV35 in Africa, which should be included in future vaccines to reduce regional disparities. Our findings also provide new perspectives on the inclusion of individual HPV genotypes in cervical cancer screening tests, and their meaningful grouping for clinical risk management, which will become increasingly important for screening populations vaccinated against the most carcinogenic HPV genotypes.

Contributors

GMC conceived the original idea for the study. FW and GMC did the literature search and review. FW and DG analysed the data. All authors had full access to the data. FW, DG, and GMC accessed and verified the data. FW wrote the first draft of the manuscript. GMC, DG, IM, and IB revised the manuscript. All authors provided input, approved the final manuscript, and had final responsibility for the decision to submit for publication. All authors have seen and approved of the final text.

Declaration of interests

We declare no competing interests.

Data sharing

Data is available upon request to the corresponding author.

Acknowledgments

We thank Nadia Akel (International Agency for Research on Cancer/World Health Organization [IARC/WHO]) for language editing support,

Teresa Lee (IARC/WHO) for support on the literature search strategy, Latifa Bouanzi (IARC/WHO) for the acquisition of library resources, Maxime Bonjour (IARC/WHO) for help on data analysis, and Vanessa Tenet (IARC/WHO) for her help on identifying individual-level data on HPV prevalence in cervical cancer cases collected by IARC/WHO. We also acknowledge Joakim Dillner and Jiangrong Wang (Center for Cervical Cancer Elimination, Karolinska Institutet, Stockholm, Sweden) who provided additional information not included in the original publications. The research leading to these results has received funding from the EU's Horizon 2020 Research and Innovation Programme (grant agreement 847845; Project RISCC). Where authors are identified as personnel of IARC/WHO, the authors alone are responsible for the views expressed in this Article and they do not necessarily represent the decisions, policy, or views of IARC/WHO. The designations used and the presentation of the material in this Article do not imply the expression of any opinion whatsoever on the part of WHO and the IARC about the legal status of any country, territory, city, or area, or of its authorities, or concerning the delimitation of its frontiers or boundaries.

References

1 International Agency for Research on Cancer. Absolute numbers of cervical cancer incidence and mortality in 2022. 2024. https://gco.iarc.fr/today/en/dataviz/bars?types=1&mode=cancer&group_populations=1&sort_by=value0&key=total&sexes=2 (accessed March 14, 2024).

2 Walboomers JM, Jacobs MV, Manos MM, et al. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J Pathol* 1999; **189:** 12–19.

3 WHO. Global strategy to accelerate the elimination of cervical cancer as a public health problem. 2020. https://www.who.int/publications/i/item/9789240014107 (accessed March 14, 2024).

4 Schiffman M, Clifford G, Buonaguro FM. Classification of weakly carcinogenic human papillomavirus types: addressing the limits of epidemiology at the borderline. *Infect Agent Cancer* 2009; **4:** 8.

5 International Human Papillomavirus Reference Center. HPV reference clones. 2023. https://www.hpvcenter.se/human_reference_clones/ (accessed March 14, 2024).

6 International Agency for Research on Cancer. Monographs on the evaluation of carcinogenic risks to humans, Vol. 100B. A review of human carcinogens: biological agents. Lyon: International Agency for Research on Cancer, 2012.

7 Guan P, Howell-Jones R, Li N, et al. Human papillomavirus types in 115,789 HPV-positive women: a meta-analysis from cervical infection to cancer. *Int J Cancer* 2012; **131:** 2349–59.

8 Li N, Franceschi S, Howell-Jones R, Snijders PJ, Clifford GM. Human papillomavirus type distribution in 30,848 invasive cervical cancers worldwide: variation by geographical region, histological type and year of publication. *Int J Cancer* 2011; **128:** 927–35.

9 Miettinen OS. Proportion of disease caused or prevented by a given exposure, trait or intervention. *Am J Epidemiol* 1974; **99:** 325–32.

10 Stan Development Team. "RStan: the R interface to Stan." R package version 2.32.3. 2023. https://mc-stan.org/ (accessed March 25, 2024).

11 Muñoz N, Bosch FX, Castellsagué X, et al. Against which human papillomavirus types shall we vaccinate and screen? The international perspective. *Int J Cancer* 2004; **111:** 278–85.

12 Hammouda D, Muñoz N, Herrero R, et al. Cervical carcinoma in Algiers, Algeria: human papillomavirus and lifestyle risk factors. *Int J Cancer* 2005; **113:** 483–89.

13 Keita N, Clifford GM, Koulibaly M, et al. HPV infection in women with and without cervical cancer in Conakry, Guinea. *Br J Cancer* 2009; **101:** 202–08.

14 Sherpa AT, Clifford GM, Vaccarella S, et al. Human papillomavirus infection in women with and without cervical cancer in Nepal. *Cancer Causes Control* 2010; **21:** 323–30.

15 Bardin A, Vaccarella S, Clifford GM, et al. Human papillomavirus infection in women with and without cervical cancer in Warsaw, Poland. *Eur J Cancer* 2008; **44:** 557–64.

16 Mpunga T, Chantal Umulisa M, Tenet V, et al. Human papillomavirus genotypes in cervical and other HPV-related anogenital cancer in Rwanda, according to HIV status. *Int J Cancer* 2020; **146:** 1514–22.

17 Wentzensen N, Schiffman M, Dunn T, et al. Multiple human papillomavirus genotype infections in cervical cancer progression in the study to understand cervical cancer early endpoints and determinants. *Int J Cancer* 2009; **125:** 2151–58.

18 Halec G, Schmitt M, Dondog B, et al. Biological activity of probable/possible high-risk human papillomavirus types in cervical cancer. *Int J Cancer* 2013; **132:** 63–71.

19 Quint W, Jenkins D, Molijn A, et al. One virus, one lesion—individual components of CIN lesions contain a specific HPV type. *J Pathol* 2012; **227:** 62–71.

20 de Sanjose S, Quint WGV, Alemany L, et al. Human papillomavirus genotype attribution in invasive cervical cancer: a retrospective cross-sectional worldwide study. *Lancet Oncol* 2010; **11:** 1048–56.

21 Chew K, Rooney PH, Cruickshank ME, Murray GI. Laser capture microdissection and PCR for analysis of human papilloma virus infection. *Methods Mol Biol* 2005; **293:** 295–300.

22 Kalantari M, Garcia-Carranca A, Morales-Vazquez CD, et al. Laser capture microdissection of cervical human papillomavirus infections: copy number of the virus in cancerous and normal tissue and heterogeneous DNA methylation. *Virology* 2009; **390:** 261–67.

23 Mix J, Saraiya M, Hallowell BD, et al. Cervical precancers and cancers attributed to HPV types by race and ethnicity: implications for vaccination, screening, and management. *J Natl Cancer Inst* 2022; **114:** 845–53.

24 Palmer M, Katanoda K, Saito E, et al. Genotype prevalence and age distribution of human papillomavirus from infection to cervical cancer in Japanese women: a systematic review and meta-analysis. *Vaccine* 2022; **40:** 5971–96.

25 Bonjour M, Charvat H, Franco EL, et al. Global estimates of expected and preventable cervical cancers among girls born between 2005 and 2014: a birth cohort analysis. *Lancet Public Health* 2021; **6:** e510–21.

26 Castle PE. The evolving definition of carcinogenic human papillomavirus. *Infect Agent Cancer* 2009; **4:** 7.

27 Wang J, Elfström KM, Lagheden C, et al. Impact of cervical screening by human papillomavirus genotype: population-based estimations. *PLoS Med* 2023; **20:** e1004304.

28 Schiffman M, Hyun N, Raine-Bennett TR, et al. A cohort study of cervical screening using partial HPV typing and cytology triage. *Int J Cancer* 2016; **139:** 2606–15.

29 Inturrisi F, de Sanjose S, Desai KT, et al. A rapid HPV typing assay to support global cervical cancer screening and risk-based management: a cross-sectional study. *Int J Cancer* 2023; **154:** 241–50.

30 Clifford GM, de Vuyst H, Tenet V, Plummer M, Tully S, Franceschi S. Effect of HIV infection on human papillomavirus types causing invasive cervical cancer in Africa. *J Acquir Immune Defic Syndr* 2016; **73:** 332–39.

31 de Martel C, Georges D, Bray F, Ferlay J, Clifford GM. Global burden of cancer attributable to infections in 2018: a worldwide incidence analysis. *Lancet Glob Health* 2020; **8:** e180–90.

32 Clifford GM, Tully S, Franceschi S. Carcinogenicity of human papillomavirus (HPV) types in HIV-positive women: a meta-analysis from HPV infection to cervical cancer. *Clin Infect Dis* 2017; **64:** 1228–35.

33 Ibrahim Khalil A, Mpunga T, Wei F, et al. Age-specific burden of cervical cancer associated with HIV: a global analysis with a focus on sub-Saharan Africa. *Int J Cancer* 2022; **150:** 761–72.