$rac{1}{2}$ |
|
|
|
| Dipon Das, PhD<sup>1,2</sup>, Shun Hirayama, MD<sup>1,2,3</sup>\*, Ling Ay<br>Brian Zhao, BS<sup>1</sup>, Vasileios Efthymiou, MD<sup>1,2</sup>, Julia M<br>Birgitta E. Michels, PhD<sup>5</sup>, Tim Waterboer, PhD<sup>5</sup>, Jern<br>Michael S. Lawrence, PhD<sup>4,6,7</sup>, Matthew G. Crowson<br> [<br>| [<br>| | Dipon Das, PhD<sup>2,2</sup>, Shun Hirayama, MD<sup>2,2,3</sup>\*, Ling Aye<sup>4\*</sup>, Michael E. Bryan, BSc<sup>2\*</sup>, Saskia Naegele, MD<sup>2</sup><br>Brian Zhao, BS<sup>1</sup>, Vasileios Efthymiou, MD<sup>1,2</sup>, Julia Mendel, BS<sup>1</sup>, Adam S. Fisch, MD, PhD<sup>2,4</sup>, Lea Kröller<sup></sup> ,<br>,<br>,<br>,<br>, Brian Zhao, BS<sup>2</sup>, Vasileios Etthymiou, MD<sup>2</sup>1<sup>2</sup>, Julia Mendel, BS<sup>2</sup>, Adam S. Fisch, MD, PhD<sup>2,7</sup>, Lea Kröller<sup>9</sup>,<br>Birgitta E. Michels, PhD<sup>5</sup>, Tim Waterboer, PhD<sup>5</sup>, Jeremy D. Richmon, MD<sup>1,2</sup>, Viktor Adalsteinsson, PhD Birgitta E. Michels, PhD°, Tim Waterboer, PhD°, Jeremy D. Richmon, MD<sup>-,</sup>", Viktor Adalsteinsson, PhD°<br>Michael S. Lawrence, PhD<sup>4,6,7</sup>, Matthew G. Crowson, MD<sup>1,2</sup>, A. John lafrate, MD, PhD<sup>4</sup>, Daniel L. Faden,<br>MD<sup>1,2,6¥</sup><br> ,<br>et

<sup>¥</sup>Corresponding Author

Michael S. Lawrence, PhD<sup>+,0,7</sup>, Matthew G. Crowson, MD<sup>+,2</sup>, A. John lafrate, MD, PhD<sup>+</sup>, Daniel L. Faden,<br>MD<sup>1,2,6¥</sup><br>\*Equal contribution<br><sup>2</sup> Mepartment of Otolaryngology-Head and Neck Surgery, Harvard Medical School, Bos MD<sup>+,2,8</sup>\*<br>\*Equal d<br>\*Corresp<br><sup>1</sup>Departi<br><sup>2</sup>Massac<br><sup>3</sup>Departi t, <sup>2</sup>Massachusetts Eye and Ear, Boston, Massachusetts

<sup>¥</sup>Corresponding Auth<br><sup>1</sup>Department of Otol<br><sup>2</sup>Massachusetts Eye<br><sup>3</sup>Department of Otor<br>Wakayama, Japan<br><sup>4</sup>Department of Path Corresponding Matter<br>
Department of Otolary<br>
Massachusetts Eye and<br>
Department of Otorhin<br>
Makayama, Japan<br>
Department of Patholo<br>
Division of Infections a  $\frac{1}{2}$   $\frac{3}{4}$   $\frac{1}{5}$  ( <sup>3</sup>Department of Otorhinolaryngology-Head and Neck<br>Wakayama, Japan<br><sup>4</sup>Department of Pathology, Massachusetts General H<br><sup>5</sup>Division of Infections and Cancer Epidemiology, Deu<br>Germany<br><sup>6</sup>Broad Institute of MIT and Harvard, C

Department of Otorhinolaryngology-Head and Neck Surgery, Wakayama Medical University, Massachusetts<br>Department of Otorhinolaryngology-Head and Neck Surgery, Wakayama Medical University,<br>Wakayama, Japan<br>Department of Pathol  $5$ Division of Infections and Cancer Epidemiology, Deutsches Krebsforschungszentrum, Heidelberg,

<sup>6</sup>Broad Institute of MIT and Harvard, Cambridge, Massachusetts

Nakayama, Japan<br>Department of Pathology, Massachusetts General Hospital, Boston, Massachusetts<br>Division of Infections and Cancer Epidemiology, Deutsches Krebsforschungszentrum, Heidelb<br>Broad Institute of MIT and Harvard, C <sup>4</sup> Department of Pa<br><sup>5</sup> Division of Infection<br><sup>5</sup> Division of Infection<br>6 Proad Institute of <sup>7</sup><br>Krantz Family Cen<br>Massachusetts Division of Infections and Cancer Epidemiology, Deutsches Krebsforschungszentrum<br>Bermany<br>Broad Institute of MIT and Harvard, Cambridge, Massachusetts<br>Krantz Family Center for Cancer Research, Massachusetts General Hospital  $\overline{ }$ 

Division of Inference Cancer Epidemiology, 2008 and Cancer Cancer Cancer Epidemiology<br>Broad Institute of MIT and Harvard, Cambridge, Massachusetts<br>Krantz Family Center for Cancer Research, Massachusetts General Hospital Ca <sup>6</sup> Broad Institute of MIT and Harvard, Cambridge, Mass<sup>7</sup> Krantz Family Center for Cancer Research, Massachus<br><sup>7</sup> Krantz Family Center for Cancer Research, Massachus<br>Massachusetts<br>Corresponding author: Daniel L. Faden, MD Krantz Family Center for Cancer Research, Massachusetts Gener<br>Massachusetts<br>Corresponding author: Daniel L. Faden, MD, Department of Otol:<br>Harvard Medical School. Division of Head and Neck Surgical Once<br>243 Charles St, Bos Massachusetts<br>Corresponding author: Daniel L. Faden, MD, Department of Otolaryngology-Head and Neck Surgery,<br>Harvard Medical School. Division of Head and Neck Surgical Oncology, Massachusetts Eye and Ear<br>143 Charles St, Bo Massachusetts<br>Corresponding<br>Harvard Medica<br>243 Charles St, (<br>|
|-<br>| Harvard Medical School. Division of Head and Neck Surgical Oncology, Massachusetts Eye and Ear Harvard Medical School. Division of Head and Neck Surgical Oncology, Massachusetts Eye and EAR<br>243 Charles St, Boston, MA 02118 (dfaden@mgb.org) 243 Charles St, Boston, MA 02118 (dfaden@mgb.org)

|
|
| ノート いちのころ しょうしょう Abstract<br>Background: HPV-associated oropharyngeal cancer (HPV+OPSCC) is the most common HPV-associated<br>cancer in the United States yet unlike cervical cancer lacks a screening test. HPV+OPSCCs are presumed<br>to start develop Fact in the United States yet unlike cervical cancer lacks a screening test. HPV+OPSCCs are presumed<br>to start developing 10-15 years prior to clinical diagnosis. Circulating tumor HPV DNA (ctHPVDNA) is a<br>sensitive and spec to start developing 10-15 years prior to clinical diagnosis. Circulating tumor HPV DNA (ctHPVDNA) is a<br>sensitive and specific biomarker for HPV+OPSCC. Taken together, blood-based screening for HPV+OPSC<br>may be feasible year

sensitive and specific biomarker for HPV+OPSCC. Taken together, blood-based screening for HPV+OPS<br>may be feasible years prior to diagnosis.<br>Methods: We developed an HPV whole genome sequencing assay, HPV-DeepSeek, with 99% sensitive and specific biomates in the specific biomarker for HPV+OPSCC. The specific biomarker for HPV wholes<br>Methods: We developed an HPV whole genome sequencing assay, HPV-DeepSeek, with 99% sensitivity<br>and specificity may be feasible years prior to angulation.<br>Methods: We developed an HPV whole g<br>and specificity at clinical diagnosis. 28 playears prior to diagnosis along with 1:1 ag<br>and an HPV serology assay.<br>Results: 22/28 (79%) of cas 「ミート」 こうしょう and specificity at clinical diagnosis. 28 plasma samples from HPV+OPSCC patients collected 1.3-10.8<br>years prior to diagnosis along with 1:1 age and gender-matched controls were run on HPV-DeepSeek<br>and an HPV serology assay

years prior to diagnosis along with 1:1 age and gender-matched controls were run on HPV-DeepSeel<br>and an HPV serology assay.<br>Results: 22/28 (79%) of cases and 0/28 controls screened positive for HPV+OPSCC with 100% detecti<br> years prior to diagnosis and a maximum server and the term of the term of the term in the top positions and an HPV serology assay.<br>Results: 22/28 (79%) of cases and 0/28 controls screened positive for HPV+OPSCC with 100% d and an HPV serology assay.<br>Results: 22/28 (79%) of cases and 0/28 controls screened positive for HPV+OPSCC with 100% detection<br>within four years of diagnosis and a maximum lead time of 7.8 years. We next applied a machine<br> || \<br>| {|
| {| within four years of diagnosis and a maximum lead time of 7.8 years. We next applied a machine<br>learning model classifying 27/28 cases (96%) with 100% detection within 10 years. Plasma-based PIK3CA<br>gene mutations, viral gen learning model classifying 27/28 cases (96%) with 100% detection within 10 years. Plasma-based<br>gene mutations, viral genome integration events and HPV serology were used to orthogonally val<br>cancer detection with 68% (19/28 gene mutations, viral genome integration events and HPV serology were used to orthogonally validate cancer detection with 68% (19/28) of the cohort having multiple cancer signals detected. Molecular fingerprinting of HPV g gene mutations, with 68% (19/28) of the cohort having multiple cancer signals detected. Molecular fingerprinting of HPV genomes was performed across patients demonstrating that each viral genome was unique, ruling out cont fingerprinting of HPV genomes was performed across patients demonstrating that each viral genome<br>was unique, ruling out contamination. In patients with tumor blocks from diagnosis (15/28), molecula<br>fingerprinting was perfo

was unique, ruling out contamination. In patients with tumor blocks from diagnosis (15/28), molecula<br>fingerprinting was performed within patients confirming the same viral genome across time.<br>Conclusions: We demonstrate ac fingerprinting was performed within patients confirming the same viral genome across time.<br>Conclusions: We demonstrate accurate blood-based detection of HPV-associated cancers with lead<br>times up to 10 years before clinical fingerprinting was performed within patients confirming the same viral genome across inter-<br>Conclusions: We demonstrate accurate blood-based detection of HPV-associated cancers wit<br>times up to 10 years before clinical canc (<br>t<br>( Conclusions: We demonstrate accurate accuration of HPV-associations: We detection of HPV-associations: We detect<br>CONA-based cancer screening.<br>CCDNA-based cancer screening. times up to 10 years before climical cancer clinical cancer diagnosis and in doing so, highlight the entriest potential of<br>ctDNA-based cancer screening. ctDNA-based cancer screening.

י<br>|
| ז |
| ( c<br>| c Introduction<br>Circulating tumor (ct) DNA approaches are being applied across nearly all cancer care contexts, including<br>screening. Numerous Multi-Cancer Early Detection (MCED) approaches are now available in both<br>research a Experimental Circuit (Circuit Circuit Circuit Circuit Circuit Circuit Circuit Circuit Circuit Circuit Circuit<br>
The paraches are now available in both<br>
research and clinical settings<sup>1-3</sup>. Continued advancements in the plat research and clinical settings<sup>1-3</sup>. Continued advancements in the platforms and computational<br>approaches available are driving lower and lower limits of cancer detection in the blood. How early in<br>development a cancer can

research and clinical settings<sup>4</sup><sup>5</sup>. Continued advancements in the platforms and computational<br>approaches available are driving lower and lower limits of cancer detection in the blood. How<br>development a cancer can be accu approaches a cancer can be accurately detected remains to be determined<sup>4</sup>.<br>
ctDNA-based early detection approaches have the potential to drastically improve cancer screening,<br>
particularly for cancers that currently do no .<br>. ri<br>. ri<br>. n いり けいらく particularly for cancers that currently do not have screening approaches. At present, only four cancer<br>types have population-level screening tests recommended, meaning that most diagnoses are in cance<br>types for which ther types have population-level screening tests recommended, meaning that most diagnoses are in cance<br>types for which there is no screening and thus tend to be later-stage disease, when the cancer has<br>reached a size to become types for which there is no screening and thus tend to be later-stage disease, when the cancer has<br>reached a size to become symptomatic<sup>5,6</sup>. Because diagnosis of later-stage cancers results in decreased<br>survival, increas there is no screening there is no secure there is no secure there is no security all there is unique survival, increased costs, and increased treatment morbidity, detection of asymptomatic, early-stage cancers creates a un reached a size to become symptomatic<sup>2,5</sup>. Because diagnosis of later-stage cancers results in decreased<br>survival, increased costs, and increased treatment morbidity, detection of asymptomatic, early-stage<br>cancers creates

cancers creates a unique opportunity for simultaneous improvement in all three of these critical<br>domains<sup>7-12</sup>.<br>HPV-associated oropharyngeal squamous cell carcinoma (HPV+OPSCC) is the most common HPV-<br>associated malignancy domains<sup>7-12</sup>.<br>HPV-associated oropharyngeal squamous cell carcinoma (HPV+OPSCC) is the most common HPV-<br>associated malignancy in the United States, yet unlike HPV-associated cervical cancer, it lacks a<br>screening test<sup>13,14</sup> ト らくりょう tonsils and thus are frequently asymptomatic, even at the time of clinical diagnosis. Instead, the most screening test<sup>13,14</sup>. Oropharyngeal infection with HPV typically occurs >20 years prior to clinical diagnosis and genomic models suggest HPV+OPSCC begins developing 10-15 years prior to diagnosis and genomic models sugges diagnosis and genomic models suggest HPV+OPSCC begins developing 10-15 years prior to diagnosis<sup>15</sup>.<br>Further, patients with HPV+OPSCC develop HPV oncoprotein-directed antibodies that have high<br>sensitivity and specificity f diagnosis and genomic models suggest HPV+OPSCC begins developing 10-15 years prior to diagnosis<sup>15</sup>.<br>Further, patients with HPV+OPSCC develop HPV oncoprotein-directed antibodies that have high<br>sensitivity and specificity f Further, patients with HPV+OPSCC at the time of diagnosis and can be detected >10 years<br>diagnosis, in some patients<sup>16-19</sup>. HPV+OPSCCs develop inside the deep crypts of the palatine and<br>tonsils and thus are frequently asym diagnosis, in some patients<sup>16-19</sup>. HPV+OPSCCs develop inside the deep crypts of the palatine and lingual tonsils and thus are frequently asymptomatic, even at the time of clinical diagnosis. Instead, the most common prese thus and the messenting symptom for HPV+OPSCC is a palpable neck mass, as HPV+OPSCCs metastasize to<br>cervical lymph nodes in nearly all cases by the time of diagnosis. Taken together, there is a substantial<br>time window in w cervical lymph nodes in nearly all cases by the time of diagnosis. Taken together, there is a substantial<br>time window in which screening early detection of HPV+OPSCC may be feasible and enable detection o<br>earlier-stage, n

time window in which screening early detection of HPV+OPSCC may be feasible and enable detection c<br>earlier-stage, non-metastatic cancers.<br>HPV+OPSCCs release fragments of the viral cancer genome, termed circulating tumor HP earlier-stage, non-metastatic cancers.<br>HPV+OPSCCs release fragments of the viral cancer genome, termed circulating tumor HPV DNA<br>(ctHPVDNA), which is detectable at diagnosis with high sensitivity and specificity and is mor earlier-stage, non-metastatic cancers. |
| (|
| c (ctHPVDNA), which is detectable at diagnosis with high sensitivity and specificity and is more rel<br>than mutation- or methylation-based ctDNA detection<sup>20,21</sup>. Importantly, while transient orophar<br>infection with HPV is com than mutation- or methylation-based ctDNA detection<sup>20,21</sup>. Importantly, while transient oropharynges<br>infection with HPV is common, it does not lead to detectable ctHPVDNA<sup>22</sup>. A small proof-of-principle<br>study has suggest infection with HPV is common, it does not lead to detectable ctHPVDNA<sup>22</sup>. A small proof-of-principle<br>study has suggested that ctHPVDNA may be detectable in asymptomatic individuals years prior to<br>clinical diagnosis of HP infection with HPV is common, it does not lead to detectable ctHPVDNA<sup>22</sup>. A small proof-of-principle<br>study has suggested that ctHPVDNA may be detectable in asymptomatic individuals years prior to<br>clinical diagnosis of HPV clinical diagnosis of HPV+OPSCC, but that existing approaches are not sensitive enough for reliable

ctHPVDNA detection in most cases<sup>23,24</sup>.<br>In this retrospective study, we evaluate the feasibility of a screening test for HPV-associated cancer<br>asymptomatic persons, aiming to determine if a multi-modal next-generation seq |
| a<br>|
|
| In this retrospective study, and the study of the fearest sequencing (NGS) liquid<br>
biopsy can reliably detect ctHPVDNA and how far before clinical diagnosis this detection can occur.<br> **Methods**<br> *Participants*<br> *Diagnostic* biopsy can reliably detect ctHPVDNA and how far before clinical diagnosis this detection can occur.

## r urticipums<br>P

biopsytem of Schematic Conductions<br>
Participants<br>
Diagnostic cohort: 5-10ml of whole blood was collected from 153 patients presenting with HPV+OP<br>
at Mass Eye and Ear in Boston, Massachusetts. HPV+OPSCC diagnosis consisted Methods<br>*Participar*<br>*Diagnosti*<br>at Mass Exconfirmat<br>or DNA po Diagnostic conort: 5-10ml of whole blood was collected from 153 patients presenting with HPV 3CC<br>at Mass Eye and Ear in Boston, Massachusetts. HPV+OPSCC diagnosis consisted of histopathologic<br>confirmation of SCC, p16 immun confirmation of SCC, p16 immunohistochemistry, and direct HPV testing with RNA in-situ hybridization<br>or DNA polymerase chain reaction (PCR). 5-10ml of whole blood was also collected from 153 patients or DNA polymerase chain reaction (PCR). 5-10ml of whole blood was also collected from 153 patients<br>
SCCC, parameter Analysis and direct HPV testing with RNA in the RNA in-situ hybridization<br>
The SNA intervals and the RNA i or DNA polymerase chain reaction (PCR). 5-10ml of whole blood was also collected from 153 patients

)<br>|<br>c |}<br>|}<br>|f presenting for routine care at Mass-Care-Line care patients had the state of patients had a history of non-HPV-associated cancers.<br>
Screening cohort: The MassGeneralBrigham biobank has prospectively collected blood samples Screening cohort: The MassGeneralBrigham b<br>from 140,000 patients presenting for care in t<br>Biobank clinical data was queried to identify p<br>after blood sample contribution, 2. had  $\ge 1$ ml<br>an HPV-associated malignancy at the 、funceric to Screening conort: The MassGeneralBrigham biobank has prospectively conceted blood samples and data<br>from 140,000 patients presenting for care in the healthcare system based in Boston, Massachusetts.<br>Biobank clinical data wa From 140,000 patients was queried to identify patients who: 1. were diagnosed with HPV+OPSCC >1 yeafter blood sample contribution, 2. had  $\ge 1$ ml of plasma available for retrieval and, 3. had no history can HPV-associated after blood sample contribution, 2. had  $\geq 1$ ml of plasma available for retrieval and, 3. had no history of<br>an HPV-associated malignancy at the time of blood sample contribution. Medical record review was<br>then undertaken an HPV-associated malignancy at the time of blood sample contribution. Medical record review was<br>then undertaken to ensure confirmation of HPV+OPSCC development and extraction of<br>clinicodemographic data, yielding a cohort then undertaken to ensure confirmation of HPV+OPSCC development and extraction of<br>clinicodemographic data, yielding a cohort of 28 patients with 1-4ml of frozen plasma. 20/28 patient<br>were eventually diagnosed within MassGe clinicodemographic data, yielding a cohort of 28 patients with 1-4ml of frozen plasma. 2<br>were eventually diagnosed within MassGeneralBrigham permitting retrieval of fresh-fro<br>embedded (FFPE) tumor tissue blocks from the ti mere eventually diagnosed within MassGeneralBrigham permitting retrieval of fresh-frozen paraffin-<br>embedded (FFPE) tumor tissue blocks from the time of diagnosis in 15 patients. 1:1 age- and sex-<br>matched controls were iden

sheared prior to sequencing. A custom hybrid capture library was designed with 1x end-to-end tiling over the full HPV genome of 43 genotypes (Supplementary Table S1). Libraries were prepared following Sample processing<br>Cell free (cf) DNA was extracted from 1-5ml of frozen plasma as previously described for all case<br>controls<sup>20</sup>. DNA was extracted from FFPE blocks after tumor identification and microdissection a<br>sheared いくいいし Sample processing<br>Cell free (cf) DNA was extracted from 1-5ml of frozen plasma as previously described for all cases and Centrols<sup>20</sup>. DNA was extracted from FFPE blocks after tumor identification and microdissection and<br>sheared prior to sequencing. A custom hybrid capture library was designed with 1x end-to-end tiling<br>over the full HPV geno controls<sup>20</sup>. DNA was extracted from FFPE blocks after tumor identification and microdissection and<br>sheared prior to sequencing. A custom hybrid capture library was designed with 1x end-to-end tilin<br>over the full HPV genom over the full HPV genome of 43 genotypes (Supplementary Table S1). Libraries were prepared follow<br>the KAPA HyperCap cfDNA Workflow v1.0 (Roche). The steps include: 1. End Repair & A-Tailing; 2. KA<br>Universal UMI Adapter Lig the KAPA HyperCap cfDNA Workflow v1.0 (Roche). The steps include: 1. End Repair & A-Tailing; 2. KAPA<br>Universal UMI Adapter Ligation; 3. Post-Ligation Cleanup; 4. Amplification using KAPA UDI primer mixes;<br>5. Cleanup using Universal UMI Adapter Ligation; 3. Post-Ligation Cleanup; 4. Amplification using KAPA UDI primer mixes;<br>5. Cleanup using KAPA HyperPure Beads; 6. Quantification by Qubit and quality check by Tapestation; 7.<br>Preparation of 5. Cleanup using KAPA HyperPure Beads; 6. Quantification by Qubit and quality check by Tapestation; 7.<br>Preparation of 4-plex DNA sample library pools; 8. Use of customized KAPA HyperCap Target Enrichment<br>Probes for hybridi Preparation of 4-plex DNA sample library pools; 8. Use of customized KAPA HyperCap Target Enrichment<br>Probes for hybridization and amplification of the target libraries; 9. Cleanup of enriched target libraries<br>using KAPA Hy Probes for hybridization and amplification of the target libraries; 9. Cleanup of enriched target libraries using KAPA HyperCapture Bead Kit; 10. Quantification and quality check by Qubit and Tapestation; 11. Pooling and n using KAPA HyperCapture Bead Kit; 10. Quantification and quality check by Qubit and Tapestation; 11.<br>Pooling and normalization of libraries; 12. Denaturation; and 13. Paired-end sequencing on an Illumina<br>platform to ~6000x Pooling and normalization of libraries; 12. Denaturation; and 13. Paired-end sequencing on an Illumina<br>platform to ~6000x coverage per sample. Optimal thresholds for sensitivity and specificity were<br>established using seria platform to ~6000x coverage per sample. Optimal thresholds for sensitivity and specificity were<br>established using serial dilutions (1ng – 0.00001ng) of HPV+OPSCC patient cfDNA into a mixture of<br>control patient cfDNA run in

### Sample analysis Sample analysis

established using serial dilutions (1ng – 0.00001ng) of HPV+OPSCC patient cfDNA into a mixture<br>control patient cfDNA run in triplicate (Supplementary Table S2).<br>Sample analysis<br>Sequencing reads underwent adapter trimming a established using serial dilutions (2ng – 0.0001) of HPV+OPSC patient cfDM-M10 a minimized using control patient cfDNA run in triplicate (Supplementary Table S2).<br>
Sample analysis<br>
Sequencing reads underwent adapter trimmi Sample analysis<br>Sequencing reads underwent adapter trimming and quality filteri<br>a combined reference genome encompassing human hg38 and 64<br>Molecular Index (UMI) identification and deduplication, HPV geno<br>an HPV reference g a combined reference genome encompassing human hg38 and 64 HPV reference genomes. After Unique<br>Molecular Index (UMI) identification and deduplication, HPV genotypes were inferred from mapping to<br>an HPV reference genome. Th Molecular Index (UMI) identification and deduplication, HPV genotypes were inferred from mapping to<br>an HPV reference genome. The number of unique HPV-mapped UMIs in each sample was used to<br>determine the ctHPVDNA read count an HPV reference genome. The number of unique HPV-mapped UMIs in each sample was used to<br>determine the ctHPVDNA read counts. For HPV16 samples with > 50% genome coverage, position-wise<br>base counts were obtained, and a majo determine the ctHPVDNA read counts. For HPV16 samples with > 50% genome coverage, position-<br>base counts were obtained, and a major haplotype was created by taking the most prevalent nucle<br>at each position. These haplotypes base counts were obtained, and a major haplotype was created by taking the most prevalent nucleotide<br>at each position. These haplotypes were fed into the multi-alignment tool MUSCLE, along with multiple<br>reference builds fo at each position. These haplotypes were fed into the multi-alignment tool MUSCLE, along with multiple<br>reference builds for HPV16 sublineages. The resulting alignment was fed into PhyML (run using TOPALi<br>v2.5), assuming a t reference builds for HPV16 sublineages. The resulting alignment was fed into PhyML (run using TOPALi<br>v2.5), assuming a transversion model, allowing for invariant sites and specifying a gamma distribution.<br>The resulting phy reference build in the resulting and subject the result of the resulting and specifying a gamma distribution.<br>The resulting phylogenetic tree enabled sublineage assignment of each sample by nearest distance to a<br>sublineage

The resulting phylogenetic tree enabled sublineage assignment of each sample by nearest distance to a<br>sublineage reference genome.<br>For each sample, the HPV genotype and sublineage were inferred, and the data was re-aligned Sublineage reference genome.<br>The reach sample, the HPV genotype and sublineage were inferred, and the data was re-aligned to the<br>appropriate sublineage reference genome. Clonal mutations were called using beftools, filtere For each sample, the HPV genom<br>appropriate sublineage reference on allele fraction (> 0.5) and reading<br>Viewer (IGV). For cross-sample<br>harbored a pattern of variants ー こうしょう しゅうしょう しゅうしょう けいしゃ しゅうしゃ しゅうしゃ しゅうしゃ しゅうしゃ しゅうしゃ appropriate sublineage reference genome. Clonal mutations were called using beftools, filtered based<br>on allele fraction (> 0.5) and read depth ( $\geq 4$ ), and then manually reviewed in Integrated Genomics<br>Viewer (IGV). For on allele fraction (> 0.5) and read depth (≥4), and then manually reviewed in Integrated Genomics<br>Viewer (IGV). For cross-sample contamination analysis, samples were considered unique if they<br>harbored a pattern of variants Viewer (IGV). For cross-sample contamination analysis, samples were considered unique if they<br>harbored a pattern of variants not present in any other sample. For tumor-plasma sample concord<br>harbored a pattern of variants n harbored a pattern of variants not present in any other sample. For tumor-plasma sample concerned unique if they consider they are considered unique if they concern analysis, sample if they consider they are considered uni harbored a pattern of variants not present in any other sample. For tumor-plasma sample concordance

)<br>|
| ج<br>| c c l c<br>| c analysis, the mutation process were identified from viral-human chimeric reads after alignment to the combined HPV and human reference genome, with high confidence assigned to events having at least two supporting reads an alignment to the combined HPV and human reference genome, with high confidence assigned to e<br>having at least two supporting reads and HPV coverage between 10% and 95% of the entire contig.<br>Single-nucleotide variants (SNVs) having at least two supporting reads and HPV coverage between 10% and 95% of the entire contig.<br>Single-nucleotide variants (SNVs) in PIK3CA were called using MuTect v1 and annotated by SnpEff. All<br>mutation calls were manua Single-nucleotide variants (SNVs) in PIK3CA were called using MuTect v1 and annotated by SnpEff. A<br>mutation calls were manually reviewed in IGV by three researchers independently (LA, SH, DD) to<br>remove low-evidence mutatio

## Machine learning

mutation calls were manually reviewed in IGV by three researchers independently (LA, SH, DD) to<br>remove low-evidence mutations and artifacts.<br>Machine learning<br>A binary machine learning task was employed to classify samples mutation calls were mutations and artifacts.<br>
Machine learning<br>
A binary machine learning task was employed to classify samples into HPV-associated cancer or no<br>
cancer. Before modeling, standard data preprocessing techniq Machine learning<br>Machine learning<br>A binary machine learning task was employed t<br>cancer. Before modeling, standard data prepro<br>normalization was performed on numeric featt<br>out 20% for testing. We trained several interpr<br>Ada Machine learning Framer, Mefore modeling, standard data preprocessing techniques were implemented. Min-max<br>normalization was performed on numeric features. We allocated 80% of the data for training and h<br>out 20% for testing. We trained sev commalization was performed on numeric features. We allocated 80% of the data for training an out 20% for testing. We trained several interpretable model architectures, including Random Fo AdaBoost, and Naive Bayes. Model normalization was performalization was performalizations including Random Forest, AdaBoost, and Naive Bayes. Model performance metrics included accuracy, precision, recall, macro F1-<br>score, specificity, negative predictive AdaBoost, and Naive Bayes. Model performance metrics included accuracy, precision, recall, macro F<br>score, specificity, negative predictive value (NPV), and the area under the receiver operating<br>characteristic curve (AUC-RO score, specificity, negative predictive value (NPV), and the area under the receiver operating<br>characteristic curve (AUC-ROC). Hyperparameter tuning was conducted using 10-fold cross-validation of<br>the training set for each score, specificity, negative, the receiver, specific curve (AUC-ROC). Hyperparameter tuning was conducted using 10-fold cross-value the training set for each model. We computed feature importance for the interpretable tree the training set for each model. We computed feature importance for the interpretable tree-based<br>models to identify influential predictors. Once the models were tuned, they were finalized, and<br>bootstrap resampling was cond models to identify influential predictors. Once the models were tuned, they were finalized, and<br>bootstrap resampling was conducted on the held-out test set to generate empirical 95% confidence<br>intervals for the models' per bootstrap resampling was conducted on the held-out test set to generate empirical 95% confide<br>intervals for the models' performance metrics. AUC-ROC curves were generated for the cross-va<br>and test sets using Stratified K-F intervals for the models' performance metrics. AUC-ROC curves were generated for the cross-validat<br>and test sets using Stratified K-Fold cross-validation with five splits. To interpret the models, we utiliz<br>the permutation intervals for the models' performance metrics. AUC-ROC curves were generated for the cross-validation<br>and test sets using Stratified K-Fold cross-validation with five splits. To interpret the models, we utilized<br>the permut and the permutation importance method. This method was applied to calculate and normalize the feature<br>importance scores for all models. We visualized the feature importance to evaluate the impact of each<br>feature on the mod the permutation inportance method. This method was applied to calculate the impact of each<br>feature on the models' predictive performance.<br>HPV serology<br>Multiplex serology testing for HPV 16, 33, 35, and 45 was performed at

importance scores for all models. The feature on the models' predictive performance.<br>
HPV serology<br>
Multiplex serology testing for HPV 16, 33, 35, and 45 was performed at the German Cancer Research<br>
Center (DKFZ) as previo Feature on the model predictive performance.<br>
HPV serology<br>
Multiplex serology testing for HPV 16, 33, 35, an<br>
Center (DKFZ) as previously described<sup>25</sup>.<br>
St*udy oversight*<br>
The MassGeneralBrigham Institutional Review B HPV serology

## Study oversight

|| 【<br>| (<br>| ( Center (DKFZ) as previously described<sup>25</sup>.<br>Study oversight<br>The MassGeneralBrigham Institutional Review Board approved this study. Patients underwent inform<br>consent. All the authors collected and analyzed the data, contribu Center (DKFZ) as previously described<br>St*udy oversight*<br>The MassGeneralBrigham Institutional Reconsent. All the authors collected and an<br>and vouch for the completeness and acc<br>of the study to the protocol. No one who The MassGeneralBrigham Institutional Review Board approved this study. Patients underwent informed The Mass Consent. All the authors collected and analyzed the data, contributed to the drafting of the manuscript,<br>and vouch for the completeness and accuracy of the data, the accuracy of the analyses, and the fidelity<br>of t and vouch for the completeness and accuracy of the data, the accuracy of the analyses, and the fidelity<br>of the study to the protocol. No one who is not an author contributed to the manuscript.<br>Results<br>HPV-DeepSeek performa

and viewer and viewer and actually of the study to the protocol. No one who is not an author contributed to the manuscript.<br> **Results**<br>
HPV-DeepSeek performance and training<br>
Following assay design, optimization, and thres **Results**<br>
HPV-DeepSeek performance and training<br>
Following assay design, optimization, and threshold setting, cfDNA from 153 patients pres<br>
HPV+OPSCC and 153 general adult population control patients were run on HPV-DeepS HPV-De<br>HPV-De<br>Followir<br>HPV+OF<br>98.7%, s<br>machine<br>had equ<br>Bayes w HPV-DeepSeek performance and training HPV+OPSCC and 153 general adult population control patients were run on HPV-DeepSeek (sensitivity<br>98.7%, specificity 98.7%) (Figure 1A, B, C). HPV-DeepSeek data output was then used to train and test<br>machine learning model 98.7%, specificity 98.7%) (Figure 1A, B, C). HPV-DeepSeek data output was then used to train and test<br>machine learning models for classifying HPV+OPSCC vs no HPV+OPSCC. The machine learning models<br>had equivalent performanc machine learning models for classifying HPV+OPSCC vs no HPV+OPSCC. The machine learning models<br>had equivalent performance with overlapping confidence intervals (Supplementary Table S3). Naïve<br>Bayes was chosen as a represen had equivalent performance with overlapping confidence intervals (Supplementary Table S3). Naïve<br>Bayes was chosen as a representative model. The AUC of this model for detecting HPV+OPSCC at<br>diagnosis was 1.0 (95% Cl 0.98-1 Bayes was chosen as a representative model. The AUC of this model for detecting HPV+OPSCC at<br>diagnosis was 1.0 (95% Cl 0.98-1.0) (Figure 1D).<br>Screening detection of HPV+OPSCC<br>28 patients who contributed a blood sample to t

Bayes was 1.0 (95% CI 0.98-1.0) (Figure 1D).<br>Bayes was a representative model. The AUC of this model is model for the AUC of this model. The AUC of the Mass General Brigham biobank more than one<br>before diagnosis of HPV+OPS Screening detection of HPV+OPSCC<br>28 patients who contributed a blood sample to t<br>before diagnosis of HPV+OPSCC and 1:1 age and ي<br>د د Screening detection of HPV+OPSCC<br>28 patients who contributed a blood sample to the MassGeneralBrigham biobank more than one year before diagnosis of HPV+OPSCC and 1:1 age and gender-matched controls were identified by database<br>Brightam biobank more than one years of Masseum biobank more than one years of the Masseum one years of the Masseum one year before diagnosis of HPV+OPSCC and  $\sigma$  and  $\sigma$  and  $\sigma$  and  $\sigma$  are identified by database by database  $\sigma$ 

|
| r<br>c ي<br>)<br>) د / search, and controls underwent library preparation and molecular barcoding with HPV-DeepSeek and<br>were loaded into a single next generation sequencing run. Using the prespecified cut-offs (Figure 1B, 0<br>22/28 (79%) of cases were loaded into a single next generation sequencing run. Using the prespecified cut-offs (Figure 1B, 22/28 (79%) of cases and 0/28 controls screened positive for HPV+OPSCC (Figure 2B). The maximum lead time from a positiv 22/28 (79%) of cases and 0/28 controls screened positive for HPV+OPSCC (Figure 2B). The maximum<br>lead time from a positive screening sample to clinical cancer diagnosis was 7.8 years. The screening<br>detection rate was 100% w lead time from a positive screening sample to clinical cancer diagnosis was 7.8 years. The screening<br>detection rate was 100% within four years of diagnosis. Samples that were collected closest to diagno<br>generally were the Lead time from rate was 100% within four years of diagnosis. Samples that were collected closest to diagnosially were the furthest from the test threshold and samples collected further from diagnosis we<br>closer to the test

detection research in the Euclidean Conservation relation respective to the test threshold, or negative, as expected (Figure 2C).<br>
We next applied the finalized machine learning model, finding improved classification accur generally were the test threshold, or negative, as expected (Figure 2C).<br>Closer to the test threshold, or negative, as expected (Figure 2C).<br>We next applied the finalized machine learning model, finding improved classifica closer to the test threshold, or negative, as expected (Figure 2C).<br>We next applied the finalized machine learning model, finding improved classification accuracy (pre-<br>specified ctHPVDNA cutoff accuracy 0.79; Naïve Bayes ハ ミュート く supplementary Table S4, Supplementary Figure S1). The machine learning model demonstrated a<br>maximum lead time from a positive screening sample to clinical cancer diagnosis of 10.3 years. The<br>screening detection rate was 10 HPV+OPSCC was that with the longest lead time, nearly 11 years. Machine learning improved the classification of the samples further from the time of diagnosis (Figure 2E).<br>Orthogonal validation of cancer detection in the b screening detection rate was 100% within 10 years of diagnosis. The only case not identified as<br>HPV+OPSCC was that with the longest lead time, nearly 11 years. Machine learning improved the<br>classification of the samples fu

HPV+OPSCC was that with the longest lead time, nearly 11 years. Machine learning improved the dassification of the samples further from the time of diagnosis (Figure 2E).<br>Orthogonal validation of cancer detection in the bl Hassification of the samples further from the time of diagnosis (Figure 2E).<br>
Orthogonal validation of cancer detection in the blood<br>
Plasma-based PIK3CA gene mutations, viral integration events, and HPV serology were used Controllect in the samples further from the time of angles in (Figure 22).<br>Controllect plasma-based PIK3CA gene mutations, viral integration events, and HPV set<br>orthogonally validate detection of cancer in the blood. PIK3C (「「( ・ 「 ・ ( ・ ) Plasma-based PIK3CA gene mutations, viral integration events, and HPV serology were used to orthogonally validate detection of cancer in the blood. PIK3CA is the most mutated gene in HPV<br>1/28 patients had pathogenic PIK3CA mutations detected by HPV-DeepSeek (Figure 3A, Suppler<br>Figure S2). Viral genome integration The Samin's and pathogenic PIK3CA mutations detected by HPV-DeepSeek (Figure 3A, Supplementary<br>Figure S2). Viral genome integration into the human genome is a hallmark of HPV-associated cancers.<br>7/28 patients had viral int 1/28 patients had viral integration into the human genome is a hallmark of HPV-associated cancers.<br>
1/28 patients had viral integration detected by HPV-DeepSeek (Figure 3A, B). HPV oncoprotein<br>
antibodies have high sensiti Figure S2, But and Viral integration detected by HPV-DeepSeek (Figure 3A, B). HPV oncoprotein<br>antibodies have high sensitivity and specificity for HPV+OPSCC at the time of diagnosis. Utilizing a 5-<br>genotype multiplexed ser 7/22 patients have high sensitivity and specificity for HPV+OPSCC at the time of diagnosis. Utilizing<br>aenotype multiplexed serology assay, 26/28 cases were tested for antibodies with 20/26 (77%)<br>having detectable antibodie antivity example multiplexed serology assay, 26/28 cases were tested for antibodies with 20/26 (77%) of candinalized serology assay, 26/28 cases were tested for antibodies with 20/26 (77%) of catalized having detectable an be multiplexed served in the multiplexed serology as a having both cther what in the V antibodies<br>(Figure 3A, 3C, Supplementary Table S5). 19/28 cases had at least two cancer diagnosis-supporting<br>features detected in the p

# Viral genome molecular fingerprinting

(Figure 3A, 3C, Supplementary Table S5). 19/28 cases had at least two cancer diagnosis-supporting<br>features detected in the plasma (Figure 3D).<br>Viral genome molecular fingerprinting<br>To ensure contamination could not account (Figure 3D).<br>
Features detected in the plasma (Figure 3D).<br>
Uiral genome molecular fingerprinting<br>
To ensure contamination could not account for screening detection we utilized viral genome molecular<br>
fingerprinting betwee Features accession in pheads (Figure 2D).<br>Viral genome molecular fingerprinting<br>To ensure contamination could not account f<br>fingerprinting between and within cases. 6/2<br>genotypes or sub-lineages (Figure 4A). As exp<br>shared fingerprinting between and within cases. 6/22 cases that screened positive displayed unique HPV<br>genotypes or sub-lineages (Figure 4A). As expected from epidemiologic studies, some viral genomes<br>shared the same HPV16 sub-li genotypes or sub-lineages (Figure 4A). As expected from epidemiologic studies, some viral genom<br>shared the same HPV16 sub-lineages, with 12/22 cases belonging to the HPV16 A1 sub-lineage, 3/<br>being D3 and 2/22 being A2. Wit genotypes or sub-lineages, with 12/22 cases belonging to the HPV16 A1 sub-lineage, 3/22<br>being D3 and 2/22 being A2. Within shared sub-lineage samples, we evaluated SNVs within each vira<br>genome demonstrating that 13/17 viru shared the same Board and 2/22 being A2. Within shared sub-lineage samples, we evaluated SNVs within each viral<br>genome demonstrating that 13/17 viruses displayed unique SNV fingerprints while the remaining fou<br>lacked SNVs,

genome demonstrating that 13/17 viruses displayed unique SNV fingerprints while the remaining four<br>lacked SNVs, precluding assignment as unique viruses (Figure 4A, Supplementary Figure S3).<br>15/28 patients were treated for genometry Figure 4A, Supplementary Figure 53).<br>15/28 patients were treated for their eventual cancer diagnosis within our healthcare system and had<br>15/28 patients were treated for their eventual cancer diagnosis within our lacked SNVs, precluding assignment as unique viruses (Figure 4A, Supplementary Figure S3).<br>15/28 patients were treated for their eventual cancer diagnosis within our healthcare system and had<br>FFPE tumor tissue blocks which ・「H」<br>「「<br>「 FFPE tumor tissue blocks which could be retrieved from diagnostic biopsy or surgery. DNA from tumor tissue was subject to HPV-DeepSeek and viral genome fingerprinting was performed to verify the same viral genome in both p tissue was subject to HPV-DeepSeek and viral genome fingerprinting was performed to verify the same<br>viral genome in both pre-diagnostic plasma and diagnostic tumor tissue. All plasma-tissue pairs shared<br>the same genome, wi The same genome in both pre-diagnostic plasma and diagnostic tumor tissue. All plasma-tissue pairs shared<br>the same genome, with matching only to the pair-mate, and to no other viral genome, confirming that<br>the same virus w the same genome, with matching only to the pair-mate, and to no other viral genome, confirming that<br>the same virus was present from screening detection to diagnosis (Figure 4B, Supplementary Figure S4)<br>Longitudinal monitor

the same virus was present from screening detection to diagnosis (Figure 4B, Supplementary Figure S4)<br>Longitudinal monitoring<br>7/28 patients had been prospectively enrolled into ctHPVDNA studies within our lab at the time o the same virus was present from screening assessor to diagnosis (Figure 4B, Supplementary Figure 4B, Supplement<br>7/28 patients had been prospectively enrolled into ctHPVDNA studies within our lab at the time of<br>cancer diagn  $\frac{1}{2}$ 7/28 patients had been prospectively enrolled into ctHPVDNA studies within our lab at the time of The cancer diagnosis unrelated to their biobank sample collection (Supplementary Figure S5). Four of the time<br>cancer diagnosis unrelated to their biobank sample collection (Supplementary Figure S5). Four of the time<br>of the cancer diagnosis unrelated to their biobank sample collection (Supplementary Figure S5). Four of these samples<br>The second to the second to

It is made available under a [CC-BY-NC-ND 4.0 International license](http://creativecommons.org/licenses/by-nc-nd/4.0/) . medRxiv preprint doi: [https://doi.org/10.1101/2024.01.04.24300841;](https://doi.org/10.1101/2024.01.04.24300841) this version posted February 2, 2024. The copyright holder for this preprint<br>(which was not certified by peer review) is the author/funder, who has granted

 |}<br>|}<br>|

patients had samples continuous global view from cancer screening detection through post-treatment monitorii<br>(Figure 4C). In one of these patients, ctHPVDNA persisted across treatments, predicting recurrence.<br>Discussion<br>Us and the methods are particle in the method in the monitor and supplementation in the methods (Figure 4C). In one of these patients, ctHPVDNA persisted across treatments, predicting recurrence.<br>
Discussion<br>
Using HPV+OPSCC (Figure 4C).<br>
Using HPV+OPSCC as a model, we demonstrate accurate detection of HPV-associated cance<br>
with 100% sensitivity and specificity within four years of clinical cancer diagnosis and a maximum leatime<br>
detection up |
|
|
|
| Usi<br>Usi<br>with 100%<br>time from c<br>detection u<br>orthogonal<br>integration<br>each patien They sensitivity and specificity within four years of clinical cancer diagnosis and a maximum lead<br>an detection to clinical diagnosis of 7.8 years. The application of machine learning extended<br>n up to ten years prior to cl with the from detection to clinical diagnosis of 7.8 years. The application of machine learning extended detection up to ten years prior to clinical diagnosis. These results were validated by examining orthogonal blood-bas detection up to ten years prior to clinical diagnosis. These results were validated by examining<br>orthogonal blood-based features that support cancer detection such as PIK3CA gene mutations, vira<br>integration events and HPV orthogonal blood-based features that support cancer detection such as PIK3CA gene mutations<br>integration events and HPV oncoprotein antibodies. Contamination was ruled out by demonstr<br>each patient harbored a unique viral ge orthogonal blood-based features and HPV oncoprotein antibodies. Contamination was ruled out by demonstrating t<br>each patient harbored a unique viral genome. In patients with tumor tissue available from diagnosis,<br>same viral integration harbored a unique viral genome. In patients with tumor tissue available from diagnosis, the same viral genome was seen in blood screening detection and tissue diagnosis. These results highlight the enormous pot Figure 1 and the enormous potential of ctDNA detection approaches for cancers that currently lack screening tests,<br>the enormous potential of ctDNA detection approaches for cancers that currently lack screening tests,<br>and f

same viral genome was potential of ctDNA detection approaches for cancers that currently lack screening tests,<br>and further demonstrate some of the most significant lead times from molecular cancer detection to<br>clinical dia and further demonstrate some of the most significant lead times from molecular cancer detection to<br>clinical diagnosis reported.<br>Virally induced cancers have been a long-standing model system for early detection studies as clinical diagnosis reported.<br>
Virally induced cancers have been a long-standing model system for early detection studies as<br>
they afford a unique ctDNA target, simplifying the search for these rare cancer fragments in the Virally induced cand<br>they afford a unique ctDNA<br>and because they induce un<br>Lo and colleagues have den<br>EBV DNA, and that this dete<br>Serologic markers of EBV in<br>associated nasopharyngeal and a unique ctDNA target, simplifying the search for these rare cancer fragments in the blood, ause they induce unique immunologic responses. In EBV-associated nasopharyngeal carcinoma<br>colleagues have demonstrated the ab and because they induce unique immunologic responses. In EBV-associated nasopharyngeal carcinoma<br>Lo and colleagues have demonstrated the ability to detect asymptomatic cancers using circulating tume<br>EBV DNA, and that this Lo and colleagues have demonstrated the ability to detect asymptomatic cancers using circulating tumo<br>EBV DNA, and that this detection leads to diagnosis of earlier stage disease and improved survival<sup>26</sup>.<br>Serologic marke EBV DNA, and that this detection leads to diagnosis of earlier stage disease and improved survival<sup>26</sup>.<br>Serologic markers of EBV infection have also been shown to be an effective screening approach for EBV-<br>associated nas EBV DNA, and that this detection leads to diagnosis of earlier stage disease and improved survival<sup>20</sup>.<br>Serologic markers of EBV infection have also been shown to be an effective screening approach for E<br>associated nasopha HPV-associated malignancy in the United States. associated nasopharyngeal carcinoma in select populations<sup>27</sup>. Importantly, cancers induced by<br>infectious agents account for 15-20% of all malignancies worldwide. HPV is the cause of 5% of a<br>malignancies, accounting for 68 infering and the counting for 684,000 new cancer diagnoses per year<sup>28-30</sup>. While Pap smear screent<br>led to a significant decline in cervical cancer incidence in high-income countries, cervical cancer re<br>the second leading malignancies, accounting for 684,000 new cancer diagnoses per year<sup>2950</sup>. While Pap smear screening has<br>led to a significant decline in cervical cancer incidence in high-income countries, cervical cancer remains<br>the second

the second leading cause of cancer death for women worldwide, and no population-level screening tests<br>are utilized for the remaining five HPV-associated cancers, including HPV+OPSCC, the most common<br>HPV-associated malignan are utilized for the remaining five HPV-associated cancers, including HPV+OPSCC, the most common<br>HPV-associated malignancy in the United States.<br>HPV+OPSCCs lead to measurable immunologic response through HPV oncoprotein an HPV-associated malignancy in the United States.<br>
HPV-associated malignancy in the United States.<br>
HPV+OPSCCs lead to measurable immunologic response through HPV oncoprotein antibodie<br>
potential screening biomarker. However HPV+OPSCCs lead to measurable immunipotential screening biomarker. However, the long<br>HPV+OPSCC development, and the non-linear naiseropositivity has limited clinical utility. Further, t<br>seropositivity in anogenital HPV-ass Il screening biomarker. However, the long and variable time between seroconversion and<br>SCC development, and the non-linear nature of the biomarker, means HPV oncoprotein<br>itivity has limited clinical utility. Further, the p POTENTIFY-OPSCC development, and the non-linear nature of the biomarker, means HPV oncoprotein<br>seropositivity has limited clinical utility. Further, the performance characteristics of HPV oncoprote<br>seropositivity in anogen seropositivity has limited clinical utility. Further, the performance characteristics of HPV oncopro<br>seropositivity in anogenital HPV-associated cancers are significantly worse<sup>31</sup>. ctHPVDNA detectic<br>emerged as a promising seropositivity in anogenital HPV-associated cancers are significantly worse<sup>31</sup>. ctHPVDNA detection has emerged as a promising biomarker for diagnosing and monitoring HPV+OPSCCs, which is being widely applied across multip applied across multiple clinical contexts<sup>20,32-34</sup>. The potential of ctHPVDNA as a screening tool for HPV-<br>associated cancers has been hampered by limitations in the sensitivity of available assays for ultra-low<br>ctDNA le applied across multiple clinical contexts<sup>20,32</sup><sup>24</sup>. The potential of ctHPVDNA as a screening tool for HPV-associated cancers has been hampered by limitations in the sensitivity of available assays for ultra-low<br>ctDNA lev ctDNA levels expected in a screening setting. However, HPV+OPSCC presents an ideal scenario for<br>studying blood-based early cancer detection considering the long natural history of HPV infection in the<br>oropharynx, and the p ctudying blood-based early cancer detection considering the long natural history of HPV infection is oropharynx, and the presumed development of HPV+OPSCC years or even decades prior to clinica diagnosis, supported by HPV studying blood-barying proced-based early mean development of HPV+OPSCC years or even decades prior to clinical<br>diagnosis, supported by HPV oncoprotein studies and genomic modeling<sup>15-19</sup>. The findings of this study,<br>demon

diagnosis, supported by HPV oncoprotein studies and genomic modeling<sup>15-19</sup>. The findings of this studemonstrating remarkable lead times from molecular detection to clinical diagnosis, fit with this presumed natural histor diagnosis, supported by HPV oncoprotein studies and genomic modeling<sup>2922</sup>. The findings of this study,<br>demonstrating remarkable lead times from molecular detection to clinical diagnosis, fit with this<br>presumed natural his presumed natural history.<br>
We evaluated the relationship between HPV serology and ctHPVDNA and found that ctHI<br>
had at least equivalent sensitivity to HPV serology overall (ctHPVDNA 79% vs Serology 77%) with<br>
for improved We evaluated the<br>had at least equivalent sen<br>for improved sensitivity of<br>years of diagnosis). ctHPVI<br>approaches for detecting c exast equivalent sensitivity to HPV serology overall (ctHPVDNA 79% vs Serology 77%) with a trend<br>oved sensitivity of ctHPVDNA closer to diagnosis (ctHPVDNA 100% vs Serology 86% within four<br>diagnosis). ctHPVDNA has several for improved sensitivity of ctHPVDNA closer to diagnosis (ctHPVDNA 100% vs Serology 86% within four<br>years of diagnosis). ctHPVDNA has several additional advantages over serologic testing, including that<br>approaches for dete years of diagnosis). ctHPVDNA has several additional advantages over serologic testing, including that<br>approaches for detecting ctHPVDNA are more widely available and reproducible, and that ctHPVDNA is<br>proaches for detecti supproaches for detecting ctHPVDNA are more widely available and reproducible, and that ctHPVDNA is<br>approaches for detecting ctHPVDNA are more widely available and reproducible, and that ctHPVDNA is approaches for detecting channels channels channels are more widely available and reproducible, and that channels  $\alpha$ 

ו<br>r<br>r<br>r こくごく こくこう a dynamics as evidenced by the longitudinal monitoring presented.<br>
Our findings raise practical and ethical questions, including how practitioners or patients shact on a positive ctHPVDNA screening test, considering these Contently, and the longitudinal monographs of the longitudinal deviation and ethical questions, includinal at on a positive ctHPVDNA screening test, considering these molicinically apparent. While our study demonstrates th positive ctHPVDNA screening test, considering these molecularly-detected cancers are likely not<br>apparent. While our study demonstrates the feasibility of blood-based detection of HPV-<br>ed cancers years prior to clinical dia clinically apparent. While our study demonstrates the feasibility of blood-based detection of HPV-<br>associated cancers years prior to clinical diagnosis, it does not attempt to determine whether early<br>detection would benefi HPV-associated malignancies leading to treatment with tolerable, low-morbidity approaches such as associated malignancies leading to treatment with tearlier detection of asymptomatic cancers would be beneficial, several questions must be addressed to prove this, including whether earlier detection leads to improved sur would be beneficial, several questions must be addressed to prove this, including whether earlier<br>detection leads to improved survival, lower cost, and decreased treatment morbidity, and whether<br>patient anxiety and the rec detection leads to improved survival, lower cost, and decreased treatment morbidity, and whethe<br>patient anxiety and the recently described "cancer patient in waiting" scenario can be adequately<br>managed to attain overall im patient anxiety and the recently described "cancer patient in waiting" scenario can be adequately<br>managed to attain overall improvement in quality of life<sup>35</sup>. Long-term, we envision early detection o<br>HPV-associated malign parameter, the overall improvement in quality of life<sup>35</sup>. Long-term, we envision early detection HPV-associated malignancies leading to treatment with tolerable, low-morbidity approaches such personalized cancer vaccines, managed to attain overall improvement in quality of life<sup>35</sup>. Long-term, we envision early detection of<br>HPV-associated malignancies leading to treatment with tolerable, low-morbidity approaches such as<br>personalized cancer personalized cancer vaccines, with the presumption that early cancers are more therapeutically<br>tractable and thus some could be ablated before ever becoming clinically apparent, while others may<br>not respond and would need personal and thus some could be ablated before ever becoming clinically apparent, while others<br>not respond and would need to be monitored molecularly until they became clinically apparent.<br>development of screening methods not respond and would need to be monitored molecularly until they became clinically apparent. The<br>development of screening methods as demonstrated here must therefore occur in concert with the<br>development of treatment and not response to fiscening methods as demonstrated here must therefore occur in concert with the development of treatment and monitoring plans and approaches to address the psychosocial dimensions of such scenarios. Further development of treatment and monitoring plans and approaches to address the psychosocial<br>dimensions of such scenarios. Further, the appropriate selection of patients and timing for screening<br>requires careful consideration dimensions of such scenarios. Further, the appropriate selection of patients and timing for screquires careful consideration and study. Multiple potential approaches have been proposed, serial screening of those at highest requires careful consideration and study. Multiple potential approaches have been proposed, includin<br>serial screening of those at highest risk for HPV-associated malignancies, such as individuals with HIV,<br>single screening required and those at highest risk for HPV-associated malignancies, such as individuals with HIV,<br>single screening timepoints for moderate risk inviduals, such as men at age 50, or population-level<br>application through the single screening timepoints for moderate risk inviduals, such as men at age 50, or population-level<br>application through the inclusion of ctHPVDNA as one marker in an MCED<sup>36</sup>. The utility, cost, and<br>population-level diagno

single screening times for moderate risk invidences, and population through the inclusion of ctHPVDNA as one marker in an MCED<sup>36</sup>. The utility, cost, and population-level diagnostic performance of these and other screenin prospective study.<br>
In conclusion, using plasma samples from asymptomatic individuals who went on to develop<br>
HPV+OPSCC and healthy population-level control patient samples, we demonstrate sensitive and<br>
specific blood-bas In conclusion<br>In conclusion<br>HPV+OPSCC and he<br>specific blood-base<br>diagnosis. These re<br>opportunity for a b<br>Acknowledgments IN EXEC and healthy population-level control patient samples, we demonstrate sensitive and<br>blood-based detection of HPV-associated cancers extending to a decade prior to clinical cancers.<br>These results highlight the enormo specific blood-based detection of HPV-associated cancers extending to a decade prior to clinical c<br>diagnosis. These results highlight the enormous potential of ctDNA screening approaches and rais<br>opportunity for a blood-ba diagnosis. These results highlight the enormous potential of ctDNA screening approaches and raise the<br>opportunity for a blood-based screening test for 5% of all cancers worldwide.<br>Acknowledgments<br>Funding Source<br>salary supp

diagnosis. The setting magnitude is the enormous potential of comportunity for a blood-based screening test for 5% of all cancers worldwide.<br> **Acknowledgments**<br>
Funding Source<br>
Funding for this work came from NIH/NIDCR R03 **Acknowledgments**<br>Funding Source<br>Funding for this work came from NIH/NIDCR R03DE030550 (Daniel L. Faden). I<br>salary support from NIH/NIDCR K23 DE029811, NIH/NIDCR R03 DE030550 and |
|
|
| Funding Source<br>Funding Source<br>Funding for this wo<br>salary support from<br>Declaration of inter<br>Daniel L. Faden has Funding Source<br>Funding for this work came from NIH/NIDCR R03DE030550 (Daniel L. Faden). Daniel L. Faden receives

### Declaration of interests

Funding for this term came from Milly fire Similar 2002 (Came Concern, Came Concern and Microsofts)<br>Salary support from NIH/NIDCR K23 DE029811, NIH/NIDCR R03 DE030550 and NIH/NCI R21 CA267152<br>Declaration of interests<br>Danie salary of the contribution of interests<br>Daniel L. Faden has received research funding or in-kind funding from Bristol-Myers Squibb, Calico,<br>Predicine, BostonGene and Neogenomics. He has received consulting fees from Merck, Daniel L. Faden has received research funding or in-kind funding from Bristol-Myers Squibb, Calico, Predicine, BostonGene and Neogenomics. He has received consulting fees from Merck, Noetic, Chry<br>Biomedical Advisors, Arcadia and Focus. None of these sources relate to the work in this manuscrity<br>Biomedical Advisors, Arcad Biomedical Advisors, Arcadia and Focus. None of these sources relate to the work in this manuscritpt.<br>
Hereck, Arcadia and Focus. None of these sources relate to the work in this manuscritpt. Biomedical Advisors, Arcadia and Focus. None of these sources relate to the work in this manuscritpt.

- ײַ<br>|<br>∂  $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ References<br>1. Chen X, Gole J, Gore A, et al. Non-invasive early detection of cancer four years before
- 1. Conventional diagnosis using a blood test. Nature communications 2020;11(1):3475. (<br>
2. Lennon AM, Buchanan AH, Kinde I, et al. Feasibility of blood testing combined with PE<br>
screen for cancer and guide intervention. Sc
- DOI: 10.1038/s41467-020-17316-z.<br>Lennon AM, Buchanan AH, Kinde I, et al. Feasibility of blood testing combined with PET-CT to<br>screen for cancer and guide intervention. Science 2020;369(6499) (In eng). DOI:<br>10.1126/science. DOI: 10.1038/s41467-020-17316-z.<br>Lennon AM, Buchanan AH, Kinde I, et al. Feasibility of blood testing combined with PET-CT to<br>screen for cancer and guide intervention. Science 2020;369(6499) (In eng). DOI:<br>10.1126/science. 2. Lennon Am, Buchanan Am, Amateury Scribtsher, Co20,369(6499) (In eng). DOI:<br>2. 10.1126/science.abb9601.<br>3. Liu MC, Oxnard GR, Klein EA, Swanton C, Seiden MV. Sensitive and specific multi-cancer<br>detection and localization screen for the guide intervention. The term of the street procept to the street of the street in the dietection and localization using methylation signatures in cell-free DNA. Annals official journal of the European Societ Liu MC, Oxnard GR, Klein E.<br>detection and localization u<br>official journal of the Europ<br>10.1016/j.annonc.2020.02.<br>Gydush G, Nguyen E, Bae JI<br>enables duplex sequencing<br>10.1038/s41551-022-00855 detection and localization using methylation signatures in cell-free DNA. Annals of oncol<br>official journal of the European Society for Medical Oncology 2020;31(6):745-759. (In en<br>10.1016/j.annonc.2020.02.011.<br>4. Gydush G,
- official journal of the European Society for Medical Oncology 2020;31(6):745-759. (In eng). D(<br>10.1016/j.annonc.2020.02.011.<br>Gydush G, Nguyen E, Bae JH, et al. Massively parallel enrichment of low-frequency alleles<br>enables 10.1016/j.annonc.2020.02.011.<br>Gydush G, Nguyen E, Bae JH, et al. Massively parallel enrichment of low-frequency alleles<br>enables duplex sequencing at low depth. Nat Biomed Eng 2022;6(3):257-266. (In eng). DOI:<br>10.1038/s4155 10.1016/j.annonc.2020.02.011.
- 4. Gydantical Gydantics Christian Christia 10.1038/s41551-022-00855-9.<br>U.S. Preventive Services Task Force, Grade A and B, Category Cancer.<br>(https://uspreventiveservicestaskforce.org/uspstf/topic\_search\_results?topic\_status=All&g<br>s%5B%5D=A&grades%5B%5D=B&category%5
- U.S. Preventive Services Task Fe<br>
(https://uspreventiveservicesta<br>
<u>s%5B%5D=A&grades%5B%5D=</u><br>
Sung H, Ferlay J, Siegel RL, et al<br>
Incidence and Mortality World<br>2021;71(3):209-249. (In eng). D<br>Hubbell E, Clarke CA, Aravanis (https://uspreventiveservicestaskforce.org/uspstf/topic\_search\_resul<br>
5%5B%5D=A&grades%5B%5D=B&category%5B%5D=15&searchterm)<br>
6. Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBO<br>
1ncidence and Morta
- (https://uspreventiveservicestaskips://webservicestaskips://webservicestaskips://webservicestaskips.htmates.com<br>Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of<br>Incidence and Mortali s%5B%5D=A&grades%5B%5D=B&category%5B%5D=15&searchterm).<br>Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of<br>Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin<br>
2021;71(3):209-249. (In eng). DOI: 10.3322/caac.21660.<br>
7. Hubbell E, Clarke CA, Aravanis AM, Berg CD. Modeled Reductions in Late-stage C 2021;71(3):209-249. (In eng). DOI: 10.3322/caac.21660.<br>Hubbell E, Clarke CA, Aravanis AM, Berg CD. Modeled Reductions in Late-stage Cance<br>Multi-Cancer Early Detection Test. Cancer Epidemiol Biomarkers Prev 2021;30(3):460<br>e Hubbell E, Clarke CA, Aravanis AM, Berg CD. Modeled Re<br>Multi-Cancer Early Detection Test. Cancer Epidemiol Bior<br>eng). DOI: 10.1158/1055-9965.Epi-20-1134.<br>Ahlquist DA. Universal cancer screening: revolutionary, r<br>2018;2:23. Multi-Cancer Early Detection Test. Cancer Epidemiol Biomarkers Prev 2021;30(3):460-468. (In<br>eng). DOI: 10.1158/1055-9965.Epi-20-1134.<br>8. Ahlquist DA. Universal cancer screening: revolutionary, rational, and realizable. NPJ
- eng). DOI: 10.1158/1055-9965.Epi-20-1134.<br>Ahlquist DA. Universal cancer screening: revolutionary, rational, and realizable. NPJ Precis Onc<br>2018;2:23. (In eng). DOI: 10.1038/s41698-018-0066-x.<br>WHO. Guide to cancer early dia
- 
- 2018;2:23. (In eng). DOI: 10.1038/s41698-018-0066-x.<br>WHO. Guide to cancer early diagnosis. World Health Organization; 2017.<br>Siegel RL, Miller KD, Wagle NS, Jemal A. Cancer statistics, 2023. CA Cancer J Clin 2023;73<br>48. (In
- 8. Ahlquist DA. Universal cancer screening: revolutionary, rational, and realizable. NPJ Precis Oncol<br>
2018;2:23. (In eng). DOI: 10.1038/s41698-018-0066-x.<br>
9. WHO. Guide to cancer early diagnosis. World Health Organizatio WHO. Guide to cancer early diagnosis. World Health O<br>Siegel RL, Miller KD, Wagle NS, Jemal A. Cancer statistic<br>48. (In eng). DOI: 10.3322/caac.21763.<br>Howlader N NA, Krapcho M, Miller D, Bishop K, Kosary<br>Lewis DR, Chen HS, Siegel RL, Miller KD, Wagle NS, Jemal A. Cancer statistics, 2023. CA Cance<br>48. (In eng). DOI: 10.3322/caac.21763.<br>11. Howlader N NA, Krapcho M, Miller D, Bishop K, Kosary CL, Yu M, Ruhl J, T.<br>Lewis DR, Chen HS, Feuer EJ, C 10. 48. (In eng). DOI: 10.3322/caac.21763.<br>
11. Howlader N NA, Krapcho M, Miller D, Bishop K, Kosary CL, Yu M, Ruhl J, Tatalovich Z, Mariotto A<br>
11. Howlader N NA, Krapcho M, Miller D, Bishop K, Kosary CL, Yu M, Ruhl J, Ta 19. Howlader N NA, Krapcho M, Miller D, B<br>Howlader N NA, Krapcho M, Miller D, B<br>Lewis DR, Chen HS, Feuer EJ, Cronin KA<br>National Cancer Institute. Bethesda, MI<br>November 2016 SEER data submission,<br>Clarke CA, Hubbell E, Kuria 11. Howis DR, Chen HS, Feuer EJ, Cronin KA (eds). SEER Cancer Statistics Review, 1975-2014.<br>
National Cancer Institute. Bethesda, MD. (https://seer.cancer.gov/csr/1975-2014/, based on<br>
November 2016 SEER data submission, p
- National Cancer Institute. Bethesda, MD. (https://seer.cancer.gov/csr/1975–2014/, base<br>November 2016 SEER data submission, posted to the SEER web site, April 2017.).<br>Clarke CA, Hubbell E, Kurian AW, Colditz GA, Hartman AR, November 2016 SEER data submission, posted to the SEER web site, April 2017.).<br>Clarke CA, Hubbell E, Kurian AW, Colditz GA, Hartman AR, Gomez SL. Projected Reductions in<br>Absolute Cancer-Related Deaths from Diagnosing Cance
- Clarke CA, Hubbell E, Kurian AW, Colditz GA, Hartman AR, Gomez SL. Projected Re<br>Absolute Cancer-Related Deaths from Diagnosing Cancers Before Metastasis, 200<br>Epidemiol Biomarkers Prev 2020;29(5):895-902. (In eng). DOI: 10. 21. Clarke Cancer-Related Deaths from Diagnosing Cancers Before Metastasis, 2006-2015. Can<br>
Epidemiol Biomarkers Prev 2020;29(5):895-902. (In eng). DOI: 10.1158/1055-9965. Epi-19-136<br>
Chaturvedi AK, Engels EA, Pfeiffer RM, Epidemiol Biomarkers Prev 2020;29(5):895-902. (In eng). DOI: 10.1158/1055-9965.Epi-19-1366.<br>Chaturvedi AK, Engels EA, Pfeiffer RM, et al. Human papillomavirus and rising oropharyngeal<br>cancer incidence in the United States. Chaturvedi AK, Engels EA, Pfeiffer RM, et al. Human papillomavirus and rising oropharyngeal<br>cancer incidence in the United States. Journal of clinical oncology : official journal of the<br>American Society of Clinical Oncolog cancer incidence in the United States. Journal of clinical oncology : official journal of the<br>American Society of Clinical Oncology 2011;29(32):4294-301. (In eng). DOI:<br>10.1200/jco.2011.36.4596.<br>2hang Y, Fakhry C, D'Souza
- American Society of Clinical Oncology 2011;29(32):4294-301. (In eng). DOI:<br>10.1200/jco.2011.36.4596.<br>Zhang Y, Fakhry C, D'Souza G. Projected Association of Human Papillomavirus Vaccinatio<br>Oropharynx Cancer Incidence in the American Society of Clinical Society of Clinical Oncology 2021;<br>
20.1200/jco.2011.36.4596.<br>
20.1200/jco.2011.36.4596.<br>
20.1897 Fakhry C, D'Souza G. Projected Association of Human Papillomaviru<br>
2001: 10.1001/jamaoncol.2021 20. Zhang Y, Fakhry C, D'Souza G. Projected Association of Human Papillomavirus Vaccination With<br>
14. Oropharynx Cancer Incidence in the US, 2020-2045. JAMA oncology 2021:e212907. (In eng).<br>
2001: 10.1001/jamaoncol.2021.29
- **Example, Fakhry C, D'Souza<br>2hang Y, Fakhry C, D'Souza<br>Oropharynx Cancer Inciden<br>DOI: 10.1001/jamaoncol.20<br>Leshchiner I, Mroz EA, Cha J<br>unobtainable premalignant<br>10.1038/s43018-023-00533<br>Busch CJ, Hoffmann AS, Via** DOI: 10.1001/jamaoncol.2021.2907.<br>Leshchiner I, Mroz EA, Cha J, et al. Inferring early genetic progression in cancers with<br>unobtainable premalignant disease. Nat Cancer 2023;4(4):550-563. (In eng). DOI:<br>10.1038/s43018-023-Leshchiner I, Mroz EA, Cha J, et al. Inf<br>unobtainable premalignant disease. I<br>10.1038/s43018-023-00533-y.<br>Busch CJ, Hoffmann AS, Viarisio D, et<br>asymptomatic individuals in the Ham 15. Leshchinable premalignant disease. Nat Cancer 2023;4(4):550-563. (In eng). DOI:<br>10.1038/s43018-023-00533-y.<br>16. Busch CJ, Hoffmann AS, Viarisio D, et al. Detection of stage I HPV-driven oropharynge<br>asymptomatic individ
- under 2023-00533-y.<br>10.1038/s43018-023-00533-y.<br>Busch CJ, Hoffmann AS, Viarisio D, et al. Detection of stage I HPV-driven orophary<br>asymptomatic individuals in the Hamburg City Health Study using HPV16 E6 serold Busch CJ, Hoffmann AS, Viarisic<br>asymptomatic individuals in the<br>wishing site individuals in the asymptomatic individuals in the Hamburg City Health Study using HPV16 E6 serology - A proof-<br>asymptomatic individuals in the Hamburg City Health Study using HPV16 E6 serology - A proof-

- E<br>E<br>C  $\frac{1}{2}$ 10.1016/j.eclinm.2022.101659.<br>Holzinger D, Wichmann G, Baboci L, et al. Sensitivity and specificity<br>E6 and other early proteins for the detection of HPV16-driven oropl<br>carcinoma. International journal of cancer Journal int 10.11.1012.1010.10101<br>Holzinger D, Wichmann G, Babo<br>E6 and other early proteins for t<br>carcinoma. International journa<br>2757. (In eng). DOI: 10.1002/ijc.<br>Kreimer AR, Ferreiro-Iglesias A, l<br>before OPSCC: findings from the<br>Eur
- 17. E6 and other early proteins for the detection of HPV16-driven oropharyngeal squamous cell<br>carcinoma. International journal of cancer Journal international du cancer 2017;140(12):2748-<br>2757. (In eng). DOI: 10.1002/ijc.3 carcinoma. International journal of cancer Journal international du cancer 2017;140(12):274<br>2757. (In eng). DOI: 10.1002/ijc.30697.<br>Kreimer AR, Ferreiro-Iglesias A, Nygard M, et al. Timing of HPV16-E6 antibody seroconversi 2757. (In eng). DOI: 10.1002/ijc.30697.<br>2757. (In eng). DOI: 10.1002/ijc.30697.<br>Kreimer AR, Ferreiro-Iglesias A, Nygard M, et al. Timing of HPV16-E6 antibody seroconversion<br>before OPSCC: findings from the HPVC3 consortium. Eremer AR, Ferreiro-Iglesias A, Nygard<br>Kreimer AR, Ferreiro-Iglesias A, Nygard<br>before OPSCC: findings from the HPVC3<br>European Society for Medical Oncology<br>10.1093/annonc/mdz138.<br>Waterboer T, Brenner N, Gallagher R, e<br>Oroph 18. Kreimer AR, Ferreiro-Iglesias A, Nygard M, et al. Timing of HPV16-E6 antibody seroconversion<br>before OPSCC: findings from the HPVC3 consortium. Annals of oncology : official journal of the<br>European Society for Medical O
- European Society for Medical Oncology 2019;30(8):1335-1343. (In eng). DOI:<br>10.1093/annonc/mdz138.<br>Waterboer T, Brenner N, Gallagher R, et al. Early Detection of Human Papillomavirus-Driven<br>Oropharyngeal Cancer Using Serolo Oropharyngeal Cancer Using Serology From the Study of Prevention of Anal Cancer. JAMA<br>oncology 2020;6(11):1806-1808. (In eng). DOI: 10.1001/jamaoncol.2020.4527.<br>Siravegna G, O'Boyle CJ, Varmeh S, et al. Cell free HPV DNA p
- 1991 Compharyngeal Cancer Using Serology From the Study of Prevention of Anal Cancer. JAMA<br>
19. Maconology 2020;6(11):1806-1808. (In eng). DOI: 10.1001/jamaoncol.2020.4527.<br>
20. Siravegna G, O'Boyle CJ, Varmeh S, et al. Ce Criptini (2020;6(11):1806-1808. (In eng). DOI: 10.1001/jamaoncol.2020.4527.<br>Siravegna G, O'Boyle CJ, Varmeh S, et al. Cell free HPV DNA provides an accurate and rapid<br>diagnosis of HPV-associated head and neck cancer. Clini Siravegna G, O'Boyle CJ, Varmeh S, et al. Cell free HPV DNA provides an accurat<br>diagnosis of HPV-associated head and neck cancer. Clinical Cancer Research<br>2021:clincanres.3151.2021. DOI: 10.1158/1078-0432.Ccr-21-3151.<br>Dame
- 2021:clincanres.3151.2021. DOI: 10.1158/1078-0432.Ccr-21-3151.<br>Damerla RR, Lee NY, You D, et al. Detection of Early Human Papillomavirus-Associated Cancers
- 20. Siravegna G, O'Boyle CJ, Varmeh S, et al. Cell free HPV DNA provides an accurate and rapid<br>diagnosis of HPV-associated head and neck cancer. Clinical Cancer Research<br>2021:clincanres.3151.2021. DOI: 10.1158/1078-0432.Cc ang. 1: Clincanres. 3151.2021. DOI: 10.1158/1078-0432.Ccr-21-3151.<br>2021: Clincanres. 3151.2021. DOI: 10.1158/1078-0432.Ccr-21-3151.<br>Damerla RR, Lee NY, You D, et al. Detection of Early Human Papillomavirus-A<br>by Liquid Biop Damerla RR, Lee NY, You D, et al. Detection of Early Human Papillon<br>by Liquid Biopsy. JCO precision oncology 2019;3 (In eng). DOI: 10.12<br>Tewari SR, D'Souza G, Troy T, et al. Association of Plasma Circulatin<br>Related Orophar 22. Tewari SR, D'Souza G, Troy T, et al. Association of Plasma Circulating Tumor HPV DNA With HPV-<br>Related Oropharynx Cancer. JAMA otolaryngology-- head & neck surgery 2022;148(5):488-489.<br>(In eng). DOI: 10.1001/jamaoto.20 Tewari SR, D'Souza G, Troy T, et al. Association of Plasma Circulating Tumor HPV DNA<br>Related Oropharynx Cancer. JAMA otolaryngology-- head & neck surgery 2022;148(5)<br>(In eng). DOI: 10.1001/jamaoto.2022.0159.<br>Rettig EM, Fad
- Tewari SR, D'Souza G, Troy T, et al. Association of Plasma Circulating Tumor HPV DNA With HPV-<br>Related Oropharynx Cancer. JAMA otolaryngology-- head & neck surgery 2022;148(5):488-489.<br>(In eng). DOI: 10.1001/jamaoto.2022.0 (In eng). DOI: 10.1001/jamaoto.2022.0159.<br>Rettig EM, Faden DL, Sandhu S, et al. Detection of Circulating Tumor Human Papillomavirus DNA<br>Before Diagnosis of HPV-Positive Head and Neck Cancer. International journal of cancer (In eng). Den Ernstein Depthman Hereneems Rettig EM, Faden DL, Sandhu S, et al. Detect<br>Before Diagnosis of HPV-Positive Head and International du cancer 2022 (In eng). DOI: 1<br>Rettig EM, Waterboer T, Sim E, et al. Relatio<br>t
- 22. Refore Diagnosis of HPV-Positive Head and Neck Cancer. International journal of cancer Journal<br>
24. Rettig EM, Waterboer T, Sim E, et al. Relationship of HPV16 E6 seropositivity with circulating<br>
24. Rettig EM, Waterbo
- International du cancer 2022 (In eng). DOI: 10.1002/ijc.33996.<br>Rettig EM, Waterboer T, Sim E, et al. Relationship of HPV16 E6 seropositivity with circulating<br>tumor tissue modified HPV16 DNA before head and neck cancer diag Rettig EM, Waterboer T, Sim E, et al. Relationship of HPV16 E6<br>tumor tissue modified HPV16 DNA before head and neck cance<br>2023;141:106417. (In eng). DOI: 10.1016/j.oraloncology.2023.1<br>Waterboer T, Sehr P, Michael KM, et al 222. Later tissue modified HPV16 DNA before head and neck cancer diagnosis. Oral oncology<br>
2023;141:106417. (In eng). DOI: 10.1016/j.oraloncology.2023.106417.<br>
25. Waterboer T, Sehr P, Michael KM, et al. Multiplex human pa 2023;141:106417. (In eng). DOI: 10.1016/j.oraloncology.2023.106417.<br>Waterboer T, Sehr P, Michael KM, et al. Multiplex human papillomavirus serology based c<br>situ-purified glutathione s-transferase fusion proteins. Clin Chem Waterboer T, Sehr P, Michael KM, et al. Multiplex human papillomavirusitu-purified glutathione s-transferase fusion proteins. Clin Chem 2005;<br>2001: 10.1373/clinchem.2005.052381.<br>Chan KCA, Woo JKS, King A, et al. Analysis o
- 22. Waterpurified glutathione s-transferase fusion proteins. Clin Chem 2005;51(10):1845-53. (In eng<br>
26. Chan KCA, Woo JKS, King A, et al. Analysis of Plasma Epstein-Barr Virus DNA to Screen for<br>
26. Nasopharyngeal Cancer. Situal field glutathion. 2005.052381.<br>
Chan KCA, Woo JKS, King A, et al. Analysis of Plasma Epstein-Barr Virus DNA to Screen for<br>
Nasopharyngeal Cancer. The New England journal of medicine 2017;377(6):513-522. (In eng).<br>
D
- Chan KCA, Woo JKS, King A, et al. Anal<br>Nasopharyngeal Cancer. The New Eng<br>DOI: 10.1056/NEJMoa1701717.<br>Li T, Li F, Guo X, et al. Anti-Epstein-Bar<br>Cancer. The New England journal of m<br>10.1056/NEJMoa2301496.<br>White MK, Pagano Nasopharyngeal Cancer. The New England journal of medicine 2017;377(6):513-522. (In e<br>
26. Chan Col: 10.1056/NEJMoa1701717.<br>
27. Li T, Li F, Guo X, et al. Anti-Epstein-Barr Virus BNLF2b for Mass Screening for Nasopharyng<br> DOI: 10.1056/NEJMoa1701717.<br>Li T, Li F, Guo X, et al. Anti-Epstein-Barr Virus BNLF2b for Mass Screening for Nasopharyngeal<br>Cancer. The New England journal of medicine 2023;389(9):808-819. (In eng). DOI:<br>10.1056/NEJMoa23014 Li T, Li F, Guo X, et al. Anti-Epste<br>Cancer. The New England journa<br>10.1056/NEJMoa2301496.<br>White MK, Pagano JS, Khalili K. V<br>molecular paradigms. Clinical mi<br>10.1128/cmr.00124-13.<br>de Martel C, Ferlay J, Franceschi
- 27. Cancer. The New England journal of medicine 2023;389(9):808-819. (In eng). DOI:<br>
27. 10.1056/NEJMoa2301496.<br>
28. White MK, Pagano JS, Khalili K. Viruses and human cancers: a long road of discovery of<br>
28. White MK, Pag 10.1056/NEJMoa2301496.<br>White MK, Pagano JS, Khalili K. Viruses and human cancers: a long road of discover<br>molecular paradigms. Clinical microbiology reviews 2014;27(3):463-81. (In eng). D<br>10.1128/cmr.00124-13.<br>de Martel C, White MK, Pagano JS, Khali<br>molecular paradigms. Clinic<br>10.1128/cmr.00124-13.<br>de Martel C, Ferlay J, Franc<br>2008: a review and synthet<br>10.1016/s1470-2045(12)70<br>Plummer M, de Martel C, V
- molecular paradigms. Clinical microbiology reviews 2014;27(3):463-81. (In eng). DOI:<br>
10.1128/cmr.00124-13.<br>
de Martel C, Ferlay J, Franceschi S, et al. Global burden of cancers attributable to infect<br>
2008: a review and s molecular paramger and microbiology reviews 2012; precession, the single exhibits of Martel C, Ferlay J, Franceschi S, et al. Global burden of cancers attributable to inference 2008: a review and synthetic analysis. The la 10.1128/cmr.00124-13.
- 2008: a review and synthetic analysis. The lancet oncology 2012;13(6):607-15. (In eng). DOI:<br>10.1016/s1470-2045(12)70137-7.<br>30. Plummer M, de Martel C, Vignat J, Ferlay J, Bray F, Franceschi S. Global burden of cancers<br>att 20.1016/s1470-2045(12)70137-7.<br>
Plummer M, de Martel C, Vignat J, Ferlay J, Bray F, Franceschi S. Global burden of cancers<br>
attributable to infections in 2012: a synthetic analysis. Lancet Glob Health 2016;4(9):e609-16<br>
en Plummer M, de Martel C, Vignat J,<br>attributable to infections in 2012:<br>eng). DOI: 10.1016/s2214-109x(16<br>Kreimer AR, Brennan P, Kuhs KAL,<br>Anogenital Cancer: A Nested Case 31. Physical Cancer: A Nested Case-Control Study in the European Prospective Investigation<br>31. Kreimer AR, Brennan P, Kuhs KAL, et al. Human Papillomavirus Antibodies and Future Risk<br>31. Anogenital Cancer: A Nested Case-Co
- and Political Distributable to inference of the first field of the synthesis. Dolitions in 2013: 10.1016/s2214-109x(16)30143-7.<br>Kreimer AR, Brennan P, Kuhs KAL, et al. Human Papillomavirus Antibodies and Future Risk of<br>Ano eng).<br>Kreimer AR, Brennan P, Kuhs KAL, et al. Hum<br>Anogenital Cancer: A Nested Case-Control St<br>1015-1016: A Mested Case-Control St Anogenital Cancer: A Nested Case-Control Study in the European Prospective Investigation In<br>
Anogenital Cancer: A Nested Case-Control Study in the European Prospective Investigation In<br>
Study in the European Prospective In Anogenital Cancer: A Nested Case-Control Study in the European Prospective Investigation Into

- c<br>t<br>i  $\ddot{a}$ 10.1200/jco.2014.57.8435.<br>10.1200/jco.2014.57.8435.<br>Berger BM, Hanna GJ, Posner MR, et al. Detection of Occult Recurrence Using Circu<br>Tissue Modified Viral HPV DNA among Patients Treated for HPV-Driven Oropharyn<br>Carcinoma. Everger BM, Hanna GJ, Posno<br>Tissue Modified Viral HPV D<br>Carcinoma. Clinical cancer r<br>Research 2022;28(19):4292<br>O'Boyle CJ, Siravegna G, Var<br>surgery for human papillom<br>10.1002/cncr.34109. Tissue Modified Viral HPV DNA among Patients Treated for HPV-Driven Oropharyngeal<br>Carcinoma. Clinical cancer research : an official journal of the American Association for Cancer<br>Research 2022;28(19):4292-4301. (In eng). D
- Carcinoma. Clinical cancer research : an official journal of the American Association for<br>Research 2022;28(19):4292-4301. (In eng). DOI: 10.1158/1078-0432.Ccr-22-0562.<br>O'Boyle CJ, Siravegna G, Varmeh S, et al. Cell-free hu Research 2022;28(19):4292-4301. (In eng). DOI: 10.1158/1078-0432.Ccr-22-0562.<br>C'Boyle CJ, Siravegna G, Varmeh S, et al. Cell-free human papillomavirus DNA kinetics after<br>surgery for human papillomavirus-associated orophary
- Research 2022; 28(19): An all Cell-free human papillomavirus DNA kine<br>Surgery for human papillomavirus-associated oropharyngeal cancer. Cancer 2022 (<br>10.1002/cncr.34109.<br>Chera BS, Kumar S, Beaty BT, et al. Rapid Clearance 33. O'Boyle CJ, Siravegna G, Varmeh S, et al. Cell-free human papillomavirus DNA kinetics after<br>surgery for human papillomavirus-associated oropharyngeal cancer. Cancer 2022 (In eng). I<br>10.1002/cncr.34109.<br>Chera BS, Kumar surgery for human papillomary of the Massima Circulating Tumor HPV<br>Chera BS, Kumar S, Beaty BT, et al. Rapid Clearance Profile of Plasma Circulating Tumor HPV<br>Type 16 DNA during Chemoradiotherapy Correlates with Disease Co 10.1002/cncr.34109.<br>
Chera BS, Kumar S, Beaty BT, et al. Rapid Clearance Profile of Plasma Circulating Tumor HPV<br>
Type 16 DNA during Chemoradiotherapy Correlates with Disease Control in HPV-Associated<br>
Oropharyngeal Cancer Type 16 DNA during Chemoradiotherapy Correlates with Disease Control in HPV-Associated<br>
Oropharyngeal Cancer. Clinical cancer research : an official journal of the American Associatio<br>
for Cancer Research 2019;25(15):4682-
- Type 16 Departmentation of the American Association Corpharyngeal Cancer. Clinical cancer research : an official journal of the American Association<br>for Cancer Research 2019;25(15):4682-4690. (In eng). DOI: 10.1158/1078-04
- For Cancer Research 2019;25(15):4682-4690. (In eng). DOI: 10.1158/1078-0432.Ccr-19-0211.<br>Timmermans S, Buchbinder M. Patients-in-waiting: Living between sickness and health in the<br>genomics era. J Health Soc Behav 2010;51(4 Timmermans S, Buchbinder M. Patients-in-waiting: Living between sickness and health in the<br>genomics era. J Health Soc Behav 2010;51(4):408-23. (In eng). DOI:<br>10.1177/0022146510386794.<br>Robbins HA, Ferreiro-Iglesias A, Water genomics era. J Health Soc Behav 2010;51(4):408-23. (In eng). DOI:<br>10.1177/0022146510386794.<br>36. Robbins HA, Ferreiro-Iglesias A, Waterboer T, et al. Absolute Risk of Oropharyngeal Cancer Aft<br>an HPV16-E6 Serology Test and genomics 201177/0022146510386794.<br>10.1177/0022146510386794.<br>Robbins HA, Ferreiro-Iglesias A, Waterboer T, et al. Absolute Risk of<br>an HPV16-E6 Serology Test and Potential Implications for Screening<br>Papillomavirus Cancer Coh 10.1177/0022146510386794. 36. An HPV16-E6 Serology Test and Potential Implications for Screening: Results From the Human<br>Papillomavirus Cancer Cohort Consortium. Journal of clinical oncology : official journal of the<br>American Society of Clinical On Papillomavirus Cancer Cohort Consortium. Journal of clinical oncology : official journal of the<br>American Society of Clinical Oncology 2022:Jco2101785. (In eng). DOI: 10.1200/jco.21.01785.<br>American Society of Clinical Oncol American Society of Clinical Oncology 2022:Jco2101785. (In eng). DOI: 10.1200/jco.21.01785.<br>American Society of Clinical Oncology 2022:Jco2101785. (In eng). DOI: 10.1200/jco.21.01785. American Society of Clinical Oncology 2022:Jco2101785. (In eng). DOI: 10.1200/jco.21.01785.

## Figure Legends

 $\frac{1}{2}$ |
|
|
| Figure 1. HPV-DeepSeek workflow, training and diagnostic performance<br>A. Schematic representation of HPV-DeepSeek workflow. B. Determination of HPV-DeepSeek threshold A. Schematic representation of HPV - DeepSeek workhow. B. Determination of HPV - DeepSeek threshold<br>using serial dilutions of HPV+OPSCC patient cfDNA input. Threshold was set at 10 unique HPV<br>genome reads and 10% HPV genom using serial the quantity of HPV+OPSCC patient cfDNA input. Threshold was set at 10 unique HPV<br>genome reads and 10% HPV genome coverage. Samples below the test threshold and considered<br>negative are in the gray shaded regio represent the quantity HPV genome coverage. Samples below the test threshold and considered<br>negative are in the gray shaded region. White shaded region contains positive samples. The x-axis<br>represents the number of HPV rea genome are in the gray shaded region. White shaded region contains positive samples. The x-axis represents the number of HPV reads and the y-axis represents HPV genome coverage as percent of genome. C. HPV-DeepSeek diagnos represents the number of HPV reads and the y-axis represents HPV genome coverage as percent o<br>genome. C. HPV-DeepSeek diagnostic performance in 153 HPV+OPSCC cases and 153 population-le<br>controls applying the pre-determined genome. C. HPV-DeepSeek diagnostic performance in 153 HPV+OPSCC cases and 153 population-level<br>controls applying the pre-determined thresholds. HPV+OPSCC cases are green, controls are red. HPV<br>genotypes are represented in genome. C. HPV-DeepSeek diagnostic performance in 153 HPV+OPSCC cases and 153 population-level<br>controls applying the pre-determined thresholds. HPV+OPSCC cases are green, controls are red. HPV<br>genotypes are represented in controls applying the pre-determined thresholds. Sensitivity 98.7%, Specificity 98.7%. D. Receiver<br>genotypes are represented in different shapes. Sensitivity 98.7%, Specificity 98.7%. D. Receiver<br>operating curves for three genotypes are represented in different shapes. Sensitivity 58.7%, Specificity 58.7%. D. Receiver<br>operating curves for three machine learning models for classifying HPV+OPSCC vs. no HPV+OPSC<br>AUC values are shown for all thr

AUC values are shown for all three models based on risk scores on the test cohort with 95% confidence<br>intervals (blue, RandomForest; green, AdaBoost; red, NaiveBayes).<br>Figure 2. Screening detection of HPV+OPSCC<br>A. Histogra intervals (blue, RandomForest; green, AdaBoost; red, NaiveBayes).<br> **Figure 2. Screening detection of HPV+OPSCC**<br> **A**. Histogram demonstrating blood sample collection timepoint in years prior to diagnosis for 28<br>
HPV+OPSCC Figure 2. Screening detection of HPV+OPSCC<br>A. Histogram demonstrating blood sample collection timepoint in y<br>HPV+OPSCC patients. B. Histogram showing results of HPV-DeepSe<br>demonstrating 22/28 (79%) samples screening positi | | |
| | |
| | Figure 2. Screening detection of HPV+OPSCC A. Histogram demonstrating blood sample concetion timepoint in years prior to diagnosis for 28<br>HPV+OPSCC patients. **B**. Histogram showing results of HPV-DeepSeek in 28 HPV+OPSCC patient is<br>demonstrating 22/28 (79%) samples HEV+OF SCC patients. B. Histogram showing results of HPV DeepSeek in 28 HPV+OF SCC patient samples<br>demonstrating 22/28 (79%) samples screening positive. Samples in green are positive for HPV+OPSCC<br>based on the pre-set cuto demonstrating 22/28 (1997) samples in red are negative. C. HPV-DeepSeek results of 56 samples (28<br>HPV+OPSCC, green color; 28 age- and gender- matched controls, red color). The colors of green<br>correspond to the time of coll based on the pre-set cutoris. Samples in red are negative. C. HTV-DeepDeek results of 50 samples (26<br>HPV+OPSCC, green color; 28 age- and gender- matched controls, red color). The colors of green<br>correspond to the time of c For the time of collection, with light green representing samples collected closer to diand the dark green representing samples collected further from diagnosis. Lighter-green samples to have higher numbers of reads and hi and the dark green representing samples collected further from diagnosis. Lighter-green samples tend<br>to have higher numbers of reads and higher coverage and darker-green samples tend to have lower<br>numbers of reads and lowe and the dark green represents of reads and higher coverage and darker-green samples tend to have lower<br>numbers of reads and lower coverage supporting increasing ease of detection closer to time of<br>diagnosis. **D**. Histogram numbers of reads and lower coverage supporting increasing ease of detection closer to time of<br>diagnosis. **D**. Histogram showing results of HPV-DeepSeek with machine learning demonstrating 27/:<br>(96%) samples screening posit diagnosis. **D**. Histogram showing results of HPV-DeepSeek with machine learning demonstrating<br>(96%) samples screening positive. **E**. Bar graph showing the comparison between HPV-DeepSee<br>HPV-DeepSeek with machine learning. diagnosis. D. Histogram showing results of HPV-DeepSeek with machine learning demonstrating 27/28<br>(96%) samples screening positive. E. Bar graph showing the comparison between HPV-DeepSeek and<br>HPV-DeepSeek with machine lea (96%) samples screening positive. E. Bar graph showing the comparison between in v-beepSeek and<br>HPV-DeepSeek with machine learning. Data is divided in tertiles based on collection time point. The<br>deep green color represent deep green color represents the percentage of cases determined positive by HPV-DeepSeek. The light<br>green color represents the cases detected with the addition of machine learning, showing improved<br>performance with machine

## Figure 3. Orthogonal validation of cancer detection in the blood

deep color represents the cases detected with the addition of machine learning, showing improved<br>performance with machine learning for samples further from the time of diagnosis.<br>Figure 3. Orthogonal validation of cancer d performance with machine learning for samples further from the time of diagnosis.<br> **Figure 3. Orthogonal validation of cancer detection in the blood**<br> **A**. Histogram representing the 28 HPV+OPSCC samples ordered by sample green, samples negative for a feature are in red, and untested samples are in beige. **B**. HPV-human<br>genome integration events detected in seven patients. **C**. Venn diagram for HPV serology and ctHPVDNA<br>results showing 20/2 || || ||
| || || ||
| || || A. Histogram representing the 28 HPV+OP 3ce samples ordered by sample concertor time in years prior<br>to diagnosis. Below, features supportive of cancer detection. Samples positive for a given feature are in<br>green, samples n green, samples negative for a feature are in red, and untested samples are in beige. **B**. HPV-human<br>genome integration events detected in seven patients. **C**. Venn diagram for HPV serology and ctHPVDNA<br>results showing 20/2 genome integration events detected in seven patients. C. Venn diagram for HPV serology and ctHPV<br>results showing 20/26 patients had HPV oncoprotein antibodies detected of which 17 also had<br>ctHPVDNA detected. D. Bar plot in genome integration events detected in seven patients. C. Venn diagram for HPV serology and cthPVDNA<br>results showing 20/26 patients had HPV oncoprotein antibodies detected of which 17 also had<br>ctHPVDNA detected. **D**. Bar pl results of the VDNA detected. D. Bar plot indicating the number of samples with 0, 1, 2 or 3 cancer diagred supporting features present. 25/28 cases had at least one cancer diagnosis supporting feature and at least two fea

ctriff verval detected. D. Bar plot indicating the number of samples with 0, 1, 2 or 3 cancer diagnosis<br>supporting features present. 25/28 cases had at least one cancer diagnosis supporting feature and 1<br>had at least two f supporting features and the status of eatures.<br>
Figure 4. Viral genome molecular fingerprinting and longitudinal cancer monitoring<br>
A. Phylogenetic tree of the 28 HPV+OPSCC samples based on the genotype, lineage and sublin **Figure 4. Viral genome mothly at A. Phylogenetic tree of the detected in plasma (left) a**<br>HPV16 sublineage. SNVs fo<br>Pairwise correlation heatn | ノ<br>ノ ( ト<br>|-Figure 4. Viral genome molecular fingerprinting and longitudinal cancer monitoring<br>A. Phylogenetic tree of the 28 HPV+OPSCC samples based on the genotype, lineage and sublineage A. Phylogenetic tree of the 28 HPV+OP See samples based on the genotype, imeage and submidige<br>detected in plasma (left) and molecular fingerprinting heatmap (right) for samples sharing the same<br>HPV16 sublineage. SNVs for e HPV16 sublineage. SNVs for each sample are shown in red, demonstrating that each virus is unique.<br>Pairwise correlation heatmap for the 15 samples that had FFPE tumor tissue blocks from diagnosis<br>The same same same same sam He vice submineage. SNVs for each sample are shown in red, demonstrating that each virus is unique. B.<br>Pairwise correlation heatmap for the 15 samples that had FFPE tumor tissue blocks from diagnosis Pairwise correlation heatmap for the 15 samples that had FFPE tumor tissue blocks from diagnosis

( )<br>1<br>( ) ミ(r<br>「!」 available, showing that plasma-tissue pairs correlate within a pair and not across pairs. The highest correlation of 1 is represented in red while no correlation or 0 is represented in white. C. Longitudinal monitoring of with surgery. Patients 1 and 3 were treated with chemoradiotherapy. In patients 1, 8, and 21 ctHPVDNA is cleared following treatment and remained zero during monitoring. In patient 3, ctHPVDNA was detected 20 months before HPV malignancy (asterix) for which they underwent surgery followed by chemoradiotherapy, ctHPVDNA with surgery. Patients 1 and 3 were treated with chemoradiotherapy. In patients 1, 8, and 21 ctHPVDNA<br>is cleared following treatment and remained zero during monitoring. In patient 3, ctHPVDNA was<br>detected 20 months before with surgery. Patients 1 and the mained zero during monitoring. In patient 3, ctHPVDNA was<br>detected 20 months before diagnosis. ctHPVDNA levels were monitored weekly during<br>chemoradiotherapy treatment with decreasing level is detected 20 months before diagnosis. ctHPVDNA levels were monitored weekly during<br>chemoradiotherapy treatment with decreasing levels, but no clearance. Following conclusion of<br>treatment, ctHPVDNA levels began increasing determoradiotherapy treatment with decreasing levels, but no clearance. Following concertatment, ctHPVDNA levels began increasing. The patient was then found to have a set HPV malignancy (asterix) for which they underwent treatment, ctHPVDNA levels began increasing. The patient was then found to have a second prin<br>HPV malignancy (asterix) for which they underwent surgery followed by chemoradiotherapy. ctH<br>cleared after this treatment but re HPV malignancy (asterix) for which they underwent surgery followed by chemoradiotherapy. ctHPVD<br>cleared after this treatment but re-elevated, indicating recurrence, which was detected by cross-<br>sectional imaging two months HPV manifolder, (astering the which they underwent surgery following manifolderapy.) Then there is the cleared after this treatment but re-elevated, indicating recurrence, which was detected by cross-sectional imaging two sectional imaging two months later.



# **Figure 1**













=<br>ខូ 10 <sup>-</sup><br>ប

# Figure 4



 $\overline{\mathbf{B}}$ 

