Quantification of Human Papillomavirus cell-free DNA from low volume blood plasma samples by digital PCR

Fabian Rosing[1], Matthias Meier[2], Lea Schroeder[1], Simon Laban[2], Thomas Hoffmann[2], Andreas Kaufmann[3], Oliver Siefer[4], Nora Würdemann[4], Jens Peter Klußmann[4], Thorsten Rieckmann[5, 6], Yvonne Alt[1], Daniel Faden [7, 8, 9], Tim Waterboer[1] and Daniela Höfler[1]

[1] Infections and Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany.
[2] Department of Otorhinolaryngology and Head & Neck Surgery, Head and Neck Cancer Center of the Comprehensive Cancer Center Ulm, University Medical Center Ulm, Ulm, Germany.
[3] Department of Gynecology, HPV Research Laboratory, Charité-Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany.
[4] Department of Otorhinolaryngology, Head and Neck Surgery, Medical Faculty, University of Cologne, Cologne, Germany.
[5] Department of Radiobiology & Radiation Oncology, University Medical Center Hamburg-Eppendorf, Martinistr. 52, 20246 Hamburg, Germany
[6] Department of Otolaryngology and Head and Neck Surgery, University Medical Center Hamburg-Eppendorf, Hamburg, Germany
[7] Department of Otolaryngology-Head and Neck Surgery, Boston Medical Center, Boston, Massachusetts.
[8] Department of Otolaryngology-Head and Neck Surgery, Harvard Medical School, Boston, Massachusetts.
[9] Mass Eye and Ear, Boston, Massachusetts.

Abstract

The incidence rate of human papillomavirus-driven oropharyngeal cancer (HPV-OPC) is increasing in many countries with high human development index. HPV cell-free DNA (cfDNA) isolated from relatively large amounts of blood plasma (typically 3-4 ml) has been successfully used for therapy surveillance of HPV-OPC patients. A currently highly discussed application of HPV-cfDNA is early detection of HPV-OPC in symptom-free individuals. This requires particularly sensitive and specific cfDNA detection methods as viral copy numbers can be very low. To study the predictive power of pre-diagnostic HPV-cfDNA, archived samples from epidemiological cohorts with limited plasma volume are an important source.

To establish a cfDNA detection workflow optimized for low plasma volumes we compared two cfDNA purification methods (MagNA Pure 96 and QIAamp ccfDNA/RNA) and two digital PCR systems (Biorad QX200 and QIAGEN QIAcuity One). Final assay validation included 65 low volume plasma samples from OPC patients with defined HPV status (34 HPV16+, 1 HPV33+, 2 HPV58+, 28 HPV-negative) stored for 2-9 years.

MagNA Pure 96 yielded a 28% higher cfDNA isolation efficiency in comparison to the QIAamp ccfDNA/RNA kit. Both digital PCR systems showed comparable analytic sensitivity (limit of detection 6-17 copies for both HPV16 and HPV33), but the higher multiplexing capacity allowed the QIAcuity to detect both HPV types in the same assay. In the validation set, the assay had a sensitivity of 80% (n=28/35) for detecting HPV16 and HPV33, and a specificity of 97% (n=29/30). In samples with ≥750 µl plasma available, the sensitivity was 85% (n=17/20), while in samples with <750 µl plasma the sensitivity was 73% (n=11/15).

Despite the expected drop in sensitivity with decreased plasma volume, the assay is sensitive and highly specific even in low volume samples, and thus suited for studies exploring HPV-cfDNA as an early HPV-OPC detection marker in low volume archival material.

Keywords

HPV, liquid biopsy, digital PCR, OPC, cfDNA, early detection

68 # 1. Introduction

69 Infection with high-risk human papillomavirus (HPV) is one of the leading causes of cancer
70 worldwide, accounting for 5 % of all cases [1]. HPV is a sexually transmitted infection and is
71 primarily associated with cancers of ano-genital sites, such as vulvar, penile, and anal
72 carcinomas, and particularly with more than 99% of all cervical carcinoma cases [1], but also
73 head and neck cancer, such as oropharyngeal cancer (OPC) [2]. Due to a strong increase of
74 HPV-driven OPC, OPC is now the most frequent HPV-driven cancer in some countries with
75 high human development index, especially in North America and Western Europe [2]. HPV-
76 driven OPC mostly arises from the epithelium of the tonsillar crypts and the base of tongue
77 and is often diagnosed after regional spread to nearby lymph nodes, making extensive
78 treatment necessary [3]. The clinical standard for determining the HPV status of OPC is p16-
79 immunohistochemistry (p16-IHC) [4], often accompanied by HPV DNA detection methods,
80 such as PCR. Despite the overall good prognosis of HPV-driven OPC, survivors often
81 experience serious morbidity after treatment, including speech impediments and dysphagia
82 [5]. Early detection of OPC may decrease the need for extensive treatment, possibly leading
83 to an improved quality of life. This is complicated by the fact that, contrary to cervical
84 carcinoma, no precancerous lesions of OPC are known to date [6]. Therefore, screening and
85 early detection (i.e. secondary prevention) are currently not feasible. Primary prevention
86 through HPV vaccination is highly effective at preventing cervical cancer and expectedly,
87 HPV-driven OPC as well, as the vaccine reduces the prevalence of oral high-risk HPV
88 infection [7]. However, due to OPC mainly affecting males, and the low vaccine uptake in
89 males, HPV vaccination is not expected to have an impact on HPV-OPC incidence rate for
90 decades to come. Hence, HPV-driven OPC incidence rates are projected to keep increasing
91 until at least the mid-2030s [8].

92 One biomarker that has recently gained a lot of attention in management of OPC is HPV cell-
93 free DNA (cfDNA) from blood plasma. Cells that undergo apoptosis or necrosis shed DNA
94 into their environment, and subsequently also the blood stream [9]. In healthy individuals,
95 plasma-derived cfDNA is mostly derived from hematopoietic cells [10], but in cancer
96 patients a significant fraction can be derived from the tumor. In cancer patients, the
97 proportion of tumor-derived cfDNA varies from less than 5 to more than 90% [9]. Overall
98 cfDNA concentrations can vary widely between patients, with typical values spanning from
99 10 to 1200 ng/ml [9]. As cfDNA is often of very low abundance and highly fragmented, with
100 a typical fragment size of ~180 bp [9], specialized protocols and assays are necessary for its
101 analysis. Detection of HPV cfDNA in blood plasma can identify OPC patients with high
102 sensitivity and specificity [11] and has been successfully used for diagnosis, prediction of

103 recurrence, and monitoring therapy response in OPC [12] [13] [14]. HPV cfDNA presence
104 after intended curative treatment is a highly specific predictor of recurrence, and recurrence
105 has been detected up to 5 months earlier by HPV cfDNA in comparison to standard imaging
106 techniques[15]. Therefore, HPV cfDNA is a promising candidate biomarker for a screening
107 approach due to being minimally invasive and a highly sensitive and specific biomarker for
108 HPV-driven malignancies [16]. Currently the standard method for analyzing cfDNA is digital
109 droplet PCR (ddPCR). The method relies on splitting the PCR into many thousand partitions,
110 usually by generating droplets in an emulsion (e.g. BioRad QX200), but nanoplate-based
111 approaches exist as well (e.g. QIAGEN QIAcuity), then referred to as digital PCR (dPCR). The
112 market for ddPCR devices is rapidly evolving, with new devices being capable of detecting
113 multiple targets at once. While the vast majority of HPV-OPC are caused by HPV16, there is
114 rising interest in detection of additional HPV types, especially in relatively high-incidence
115 countries [14]. In plasma samples at OPC diagnosis, the sensitivity of ddPCR has been
116 reported to range from 70% [17] to more than 98% [14]. ddPCR is more sensitive than qPCR
117 in HPV16-cfDNA detection from OPC and comes with lower costs than sequencing based
118 approaches, even though the latter have been reported to reach even higher sensitivity than
119 ddPCR [17]. Due to the fragmented nature of cfDNA in blood plasma, it is of high importance
120 for any PCR-based method to use very small amplicon sizes to maximize sensitivity of the
121 assay. Most studies have focused on application of the method for therapy surveillance or
122 diagnosis, where 3 ml of plasma or more are relatively easily available for analysis [13] [14]
123 [18]. To study the performance of HPV cfDNA as a predictive biomarker of HPV-OPC it is
124 necessary to analyze existing samples from biobanks or epidemiological cohorts, where
125 often only small amounts of plasma are available.

126 In this setting, antibodies against HPV16 oncoproteins, especially E6, have been shown to be
127 highly specific and sensitive predictive biomarkers of OPC [19]. Seroconversion has been
128 observed to occur up to 10 or more years before diagnosis [20]. HPV16 oncoprotein
129 antibody patterns have been successfully used to identify individuals at risk for development
130 of HPV-OPC in a prospective manner before onset of symptoms [21] [22]. However, due to
131 the long lead time, repeated in-depth exams by otorhinolaryngologists are necessary over
132 time for early detection of OPC, putting a considerable burden on both the patients and the
133 healthcare system. Therefore, a minimally invasive biomarker for HPV-OPC is necessary in
134 order to correctly identify people at imminent risk of developing HPV-OPC.

135 The aim of this study was to establish an optimized protocol for the purification and
136 quantification of HPV-cfDNA from small volumes of blood plasma (i.e. below 1 ml). While
137 nucleic acid purification efficiency is usually not a key determinant of cfDNA detection when

138 routine amounts of blood plasma (typically 3-4 ml) are available, it becomes more important
139 when only limited amounts of plasma are available. Blood plasma samples from volunteers
140 were used to compare cfDNA purification methods. On two different digital PCR platforms
141 (BioRad QX200 and QIAcuity One) an assay targeting HPV16, HPV33 and the human
142 reference gene beta-globin (BG) was developed to detect HPV-cfDNA. After determining
143 analytic sensitivity, DNA from cervical swab samples was analyzed to ensure HPV type
144 specificity of the assay. The 89 available liquid biosy samples from OPC patients were split
145 into a training set of 24 samples and a validation set of 65 samples. The training set was
146 used to optimize the dPCR assay, and final sensitivity and specificity were determined in the
147 validation set.

148 # 2. Material and Methods

149 ## 2.1. **OPC patient inclusion and testing**

150 Sodium citrate plasma samples from 86 OPC patients from Ulm University Hospital were
151 collected (tbl. 2.1. , Ethical approval 90/15, Ulm University Ethics Commission). Tumor HPV
152 status was determined by p16-IHC [4] and HPV DNA detection by GP5+/6+ PCR [23] and
153 subsequent sequencing of PCR products for HPV typing. OPC were classified as HPV-driven
154 by a pathologist if the tumor section was both HPV-DNA and p16-IHC positive. Tumors
155 which tested negative for either p16 or HPV-DNA (or both) were considered to be HPV-
156 negative. Three patients in the validation cohort had one blood sample taken at initial
157 diagnosis and one at recurrence with distant metastasis. Inclusion criteria for the study were
158 the diagnosis of a tumor in the oropharynx with complete tumor HPV-DNA and p16-IHC
159 data. Samples were assigned to the training and validation sets by the treating physician
160 who was not directly involved in the experiments to achieve balanced distributions of HPV-
161 driven and HPV-negative tumors in both sets. The two sets did not differ significantly when
162 examining tumor stage, nodal stage and other characteristics (tbl. 2.1). dPCR experiments
163 were conducted by personnel blinded to tumor HPV status. Available plasma volumes from
164 24 OPC patients (12 HPV16-driven, 2 HPV33-driven, 10 HPV-negative) in the training set
165 ranged from 0.38 to 2.9 ml (mean volume 1.14 ml). The samples were collected between
166 2013 and 2020 (mean storage time since blood draw 5.5 years). In the validation set of 65
167 plasma samples from 62 HPV-OPC patients (33 HPV16-driven, 2 HPV58-driven, 1 HPV33-
168 driven, 26 HPV-negative) available plasma volumes varied between 0.37 and 3 ml (mean
169 volume 1.01 ml).

170

171 **Table 2.1: Patient characteristics of the training and validation sets.**

Characteristic	**Training set** (n=24)	**Validation set** (n=62)[1]	**p-value**
HPV type (DNA)			0.38[2]
16	12 (50%)	33 (53%)	
33	2 (8.3%)	1 (1.6%)	
58	0 (0%)	2 (3.2%)	
HPV-negative[3]	10 (42%)	26 (42%)	
p16-IHC			0.89[2]
1	14 (58%)	39 (63%)	
0	10 (42%)	23 (37%)	
Age at diagnosis[4]	58 (54, 63)	61 (56, 67)[5]	0.13[6]
Sex			0.24[2]
male	21 (88%)	45 (73%)	
female	3 (13%)	17 (27%)	
OPC subsite			0.24[2]
tonsil	14 (58%)	48 (77%)	
base of tongue	9 (38%)	11 (18%)	
soft palate	1 (4.2%)	2 (3.2%)	
lateral pharyngeal wall	0 (0%)	1 (1.6%)	
T stage			0.81[2]
T1/2	13 (54%)	32 (52%)	
T3/4	11 (46%)	29 (47%)	
Unknown	0 (0%)	1 (1.6%)	
N stage			0.56[2]
N0/1	15 (63%)	44 (71%)	
N2/3	9 (38%)	17 (27%)	

Unknown	0 (0%)	1 (1.6%)	
Stage[7]			0.77[2]
Stage 1/2	14 (58%)	39 (63%)	
Stage 3/4ab	10 (42%)	21 (34%)	
Stage 4/4c	0 (0%)	1 (1.6%)	
Unknown	0 (0%)	1 (1.6%)	

N (%), [1]For 3 patients from the validation set each 2 samples were available, [2]Pearson's Chi-squared test, [3]high-risk HPV negative, [4]Median (IQR), [5]Age missing for one patient,[6]Wilcoxon rank sum test, [7]AJCC TNM8

2.2. **Healthy donor plasma sample collection**

Healthy volunteers were recruited in Heidelberg, Germany, and both EDTA and sodium citrate blood plasma samples were collected by medically trained personnel at the German Cancer Research Center (DKFZ) (Ethical approval S-031/2022, University Heidelberg Ethics Commission). Anonymized blood samples were processed within 1 h after collection by centrifugation at 600 x g for 10 minutes. The plasma phase was separated and centrifuged again at 1200 x g for 10 minutes. The resulting plasma was split into 1 ml aliquots and stored at -80°C. For each experiment, a minimum of 5 plasma sample aliquots were pooled and thoroughly mixed before proceeding.

2.3. **cfDNA isolation and purification efficiency assessment**

Before nucleic acid isolation, plasma samples were equilibrated to room temperature and centrifuged at 3000 x g for 5 minutes. Automated cfDNA purification was performed using the *DNA and Viral NA Large Volume* kit for the MagNA Pure 96 (Roche Diagnostics, Mannheim, Germany). cfDNA was purified from EDTA plasma using the *cfNA ss 2000* or *cfNA ss 4000* protocol, depending on the available plasma volume, and the *Viral NA Universal 500* or *Viral NA Universal 1000* protocol for sodium citrate plasma, as recommended by the manufacturer. The sample volume was assessed based on the sample weight and then adjusted to the necessary input volume for each protocol by adding PBS (e.g. 1500 µl plasma plus 500 µl of PBS for the *cfNA ss 2000* protocol), as recommended by the manufacturer. Purified DNA was eluted in 50 µl elution buffer.

195 The isolation efficiency of the MagNA Pure 96 protocols was compared to the *QIAamp*
196 *ccfDNA/RNA* kit (QIAGEN, Hilden, Germany) to identify the ideal protocol for cfDNA
197 isolation. Isolation by the *QIAamp* kit was conducted according to the manufacturer's
198 instructions and 20 µl of RNAse-free water was used in the final elution step. To identify the
199 optimal DNA purification method, HPV16 DNA was spiked into blood plasma from healthy
200 volunteers (spike-in experiments). 10^4 copies of a 104 bp HPV16 E6 DNA fragment
201 quantified by Agilent Bioanalyzer (Agilent High Sensitivity DNA Kit) were added to pooled
202 plasma samples from healthy volunteers immediately prior to cfDNA isolation. The pooled
203 plasma samples were then thoroughly mixed and split into several aliquots to conduct
204 cfDNA isolation and quantification in multiple replicates across the different kits. For each
205 replicate the same volume of plasma was used, and adjusted with PBS to the minimum
206 input volume of the kit, as described above. As a reference, the original material used for
207 spike-in was separately quantified in each QIAcuity dPCR run. Isolation efficiency was
208 calculated as the number of copies of spike-in DNA detected after isolation divided by the
209 number of copies detected in the original material used for spike-in.

210 2.4. **Testing HPV type specificity on cervical swab samples**

211 Cervical swabs were collected at a colposcopy clinic in Berlin (MVZ) and stored in PreservCyt
212 buffer at 4°C [24]. Samples used were HPV negative (n=21) or positive for HPV16 (n=20),
213 HPV33 (n=11), both HPV16 and HPV 33 (n=2), HPV58 (n=8), HPV31 (n=19), HPV35 (n=2), or
214 HPV52 (n=9) as determined by multiplex papillomavirus genotyping (MPG) assay [25].
215 Samples were vortexed vigorously for 20 seconds and 250 µl were used for DNA
216 purification. PBS was used to adjust the volume to the required input volume of the kit
217 used. DNA was isolated using the MagNA Pure 96 *Pathogen Universal 500* protocol, using 50
218 µl of elution buffer. 1 µl of the eluate was used for PCR.

219 2.5. **DNA quantification by dPCR and ddPCR**

220 Isolated DNA was analyzed using both the QX200 ddPCR (BioRad) and the QIAcuity dPCR
221 (QIAGEN). Both digital PCR assays quantified HPV16 E6 and the human reference gene beta-
222 globin (BG) using specific primers and probes (duplex assay) [26]. On the QIAcuity One,
223 which comes with multiple fluorescence channels to quantify up to five different targets,
224 the assay was complemented by primer-probe sets specific for HPV33 E6 (forward primer:
225 CCAAGCATTGGAGACAACTATAC, probe: TEX-CAACATTGAACTACAGTGCGTGGAATG-BHQ2,
226 reverse primer: AATCATATACCTCAGAT CGTTGC) and HPV16 E7 (forward primer:
227 AAGCAGAACCGGACAGAG, probe: HEX-TGTTGCAAGTGTGACTCTACGCTTCG-BHQ1, reverse
228 primer: TCTACGTGTGTGCTTTGTACG), thus targeting four different targets (four-plex assay).
229 These primers were designed by identifying the most conserved regions in 29 HPV33 E6 and

230 638 HPV16 E7 sequences available on NCBI GenBank. Criteria for primer and probe design
231 included suitable melting temperature (ca. 60°C) as calculated by the nearest-neighbour
232 method according to Allawi et al. (1998) [27], a maximum GC content of 70%, and minimal
233 self- and cross-complementarity to prevent the formation of primer-dimers. ddPCR on the
234 QX200 system was conducted in 40 cycles of 94°C denaturation for 30 seconds and
235 annealing and elongation at 58°C for 1 minute. dPCR on the QIAcuity One was conducted in
236 40 cycles of 95°C denaturation for 10 seconds and 60°C annealing and elongation for 30
237 seconds. For the comparison of the two PCR platforms using the training set an input of 10
238 µl of cfDNA was added per reaction, while in the validation set, where dPCR was run on
239 QIAcuity alone, the volume was 20 µl per reaction. As positive controls, and for dilution
240 series, plasmids carrying the full HPV16 and HPV33 genomes and BG were used [28,29]. The
241 positive control contained 10^4 copies of HPV16, HPV33 and human BG. Dilution series
242 spanned from $1.1x10^5$ to 1.9 copies of those templates in 3-fold dilution steps for
243 determining the analytic sensitivity of the assay.

244 2.6. **Statistical analysis**

245 Initial analysis of the QIAcuity data was conducted using the QIAcuity Software Suite V 2.2.
246 Results exported from the software were further analyzed in R version 4.3.1 [30]. Packages
247 from the tidyverse family of R-packages [31], epiR [32] and caret [33] were used. In the
248 assessment of analytic sensitivity, the limit of detection (LoD) was defined as the lowest
249 input copy number that yielded a positive result in both duplicates. In the cervix, HPV
250 infection is common, but extremely low viral loads are most likely clinically irrelevant and
251 may not have been detected by reference tests. Therefore, in cervical swab samples all
252 samples were only considered HPV-positive if a minimum of 3 partitions was positive for any
253 target. As even in blood plasma samples from patients with HPV-driven malignancies low
254 amounts of HPV DNA are expected, in the training set of plasma samples any sample with
255 any positive partition was classified as HPV-positive. Based on the results of the training set,
256 cfDNA samples were only classified as HPV-positive when both replicates were HPV cfDNA-
257 positive in the validation set of plasma samples to ensure sufficient specificity of the assay.
258 Clinical sensitivity was defined as the number of samples from patients with HPV16 or
259 HPV33-driven tumors that tested positive for HPV cfDNA divided by the number of all
260 patients with HPV16 or HPV33-driven tumors. Clinical specificity was defined as the number
261 of samples from HPV16 and HPV33-negative patients which tested negative for HPV cfDNA
262 divided by the total number of HPV16 and HPV33-negative patients. For sensitivity and
263 specificity the Clopper-Pearson confidence interval (CI) was computed.

3. Results

3.1. **Comparison of PCR platforms**

3.1.1. **Analytic sensitivity and HPV type specificity**

Analytic sensitivity of QX200 ddPCR and QIAcuity dPCR were compared by applying a three-fold serial dilution of the corresponding template DNA in duplicates. Since the QX200 ddPCR assay quantifies BG and HPV16 E6 only, these two targets were used for a direct assay comparison. The QX200 ddPCR had a LoD of 17 copies for both HPV16 E6 and BG (fig. 3.1A). The QIAcuity dPCR (fig. 3.1A and B) had a lower LoD of 6 copies for HPV16 E6, HPV16 E7 and BG, and a LoD of 17 copies for HPV33 E6.

To assess the specificity of the assay on both PCR platforms, DNA from 92 cervical swab samples with known HPV infection status was tested. Any sample with 3 or more positive partitions was counted as positive for the given target. In the QX200 ddPCR, HPV16 E6 was 100% specific (n=77/77, fig. 3.1C). Sensitivity was 95.5% (n=21/22 detected), based on one false negative sample that was defined as HPV16 and HPV33 positive by the reference method and yielded less than 3 HPV16 positive partitions. In QIAcuity dPCR, HPV16 E6 was 100% type-specific and 100% sensitive for both HPV16 single infections (n=20/20) and co-infections with HPV33 (n=2/2, fig. 3.1C). However, 12 non-HPV16 cases (out of 77) showed one to two partitions positive for HPV16 E6. HPV16 E6 quantification by the two methods was highly correlated (Pearson's r = 0.97) and virtually identical (slope of regression model = 0.93).

In QIAcuity dPCR, HPV16 E7 and HPV33 E6 were quantified in the same reaction. HPV16 E7 had a sensitivity of 100% for HPV16 single infections (n=20) and double infections with other HPV types (n=2) (fig. 3.1D). HPV33 E6 had a sensitivity of 100% for HPV 33 single infections (n=11) and double infections (n=2). While specificity for HPV16 E7 was 100%, one sample with an HPV16 infection was tested positive for HPV33, however with only 7 positive partitions, which is the lowest number detected in any sample. Consequently, the specificity of the HPV33 E6 assay was 98.8% (n=78/79). Overall, no HPV DNA was detected in any sample that was previously tested negative for HPV DNA or for HPV types other than HPV16 and HPV33.

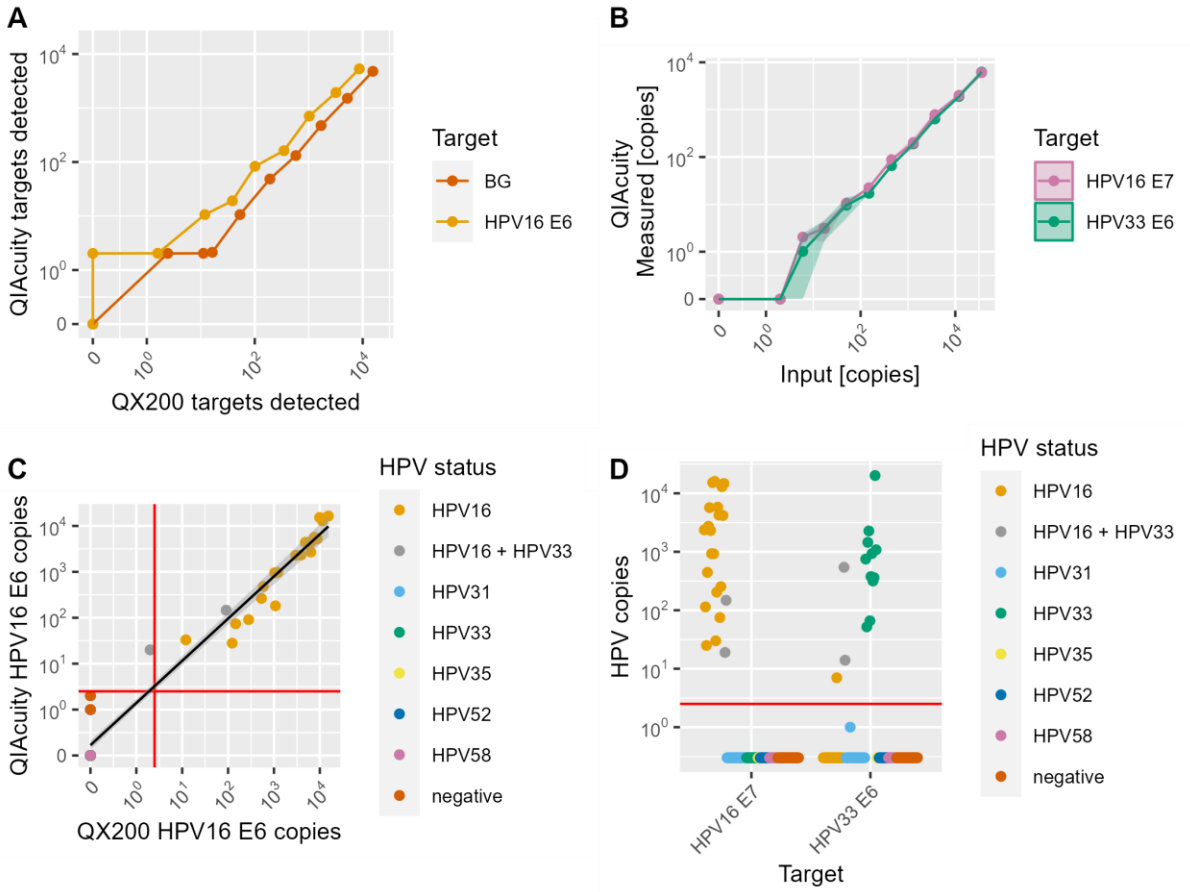


Figure 3.1: (A) Quantification of a dilution series of HPV16 E6 and beta-globin (BG) template DNA on a QX200 ddPCR system (x-axis) compared to QIAcuity dPCR (y-axis). Values shown are the minimum of duplicates. (B) Quantification of a dilution series of HPV16 and HPV33 DNA on QIAcuity dPCR by primers targeting HPV16 E7 and HPV33 E6. Values shown are the minimum of duplicates, the shaded area indicates the standard deviation. (C) Copy numbers of HPV16 E6 detected by QX200 ddPCR (y-axis) and QIAcuity dPCR (x-axis) in cervical swab samples with known HPV status. The red lines indicate the cut-off of 3 copies to distinguish positive and negative samples. The black line is a regression line fitted to the data (y = 0.11+0.93x, r=0.97). Not all colors visible due to data overlay at zero copies. (D) Copy numbers of HPV16 E7 and HPV33 E6 detected by QIAcuity dPCR in cervix samples with known HPV status. The red line indicates the cut-off of 3 copies to distinguish positive and negative samples.

3.1.2. **cfDNA analysis in a training set of 24 OPC patients**

24 plasma samples from OPC patients (12 HPV16-driven, 2 HPV33-driven, 10 HPV-negative) were analyzed to compare the QX200 ddPCR targeting HPV16 E6 and BG, and the QIAcuity dPCR targeting HPV16 E6 and BG, and HPV16 E7 and HPV33 E6 in addition. All samples were tested in duplicates.

When tested in ddPCR on the QX200 system, all samples were positive for BG in both replicates with a mean of 6,022 copies/ml detected (range 371 – 30,678 copies/ml), indicating sufficient DNA amount and quality. In total, 13 out of 24 samples had any positive

314 droplets for HPV16 E6 (range 3.33 – 8,750 copies/ml, fig. 3.2A). Three of these were false
315 positive with low HPV16 E6 copies (3.33 – 11 copies/ml), yielding a specificity of 75% (9/12,
316 95% CI: 43 – 95%). Two of these were positive for HPV33, but not HPV16, in tumor tissue.
317 The assay had a sensitivity of 83% for HPV16-driven OPC cases (10/12, 95% CI: 52 – 98%). If
318 only samples that were positive for HPV16 in both replicates were counted as HPV16
319 positive, the sensitivity decreased to 58% (7/12, 95% CI: 28 – 62%) and specificity increased
320 to 92% (11/12, 95% CI: 62 - 100%) (fig. 3.2B). The lowest concentration that was
321 reproducibly tested positive was 11 copies/ml.

322 When the assay was run with the extended dPCR target panel on the QIAcuity One, a mean
323 concentration of 6,235 copies/ml of BG per ml was detected (201 – 55,220 copies/ml),
324 indicating sufficient DNA amount and quality in all samples. The assay had an overall
325 sensitivity of 71% (10/14, 95% CI: 42 - 92%) for HPV16 (E6 and/or E7) and HPV33, and a
326 specificity of 100% (10/10, 95% CI: 69 - 100%) (fig. 3.2C). HPV33 E6 had a sensitivity of 100%
327 (2/2, 95% CI: 16 - 100%) for HPV33-driven OPC cases and a specificity of 100% (22/22, 95%
328 CI: 91 – 100%). HPV16 E6 had a sensitivity of 50% (6/12, 95% CI: 21 – 79%) and HPV16 E7 of
329 66% (8/12, 95% CI: 35 – 90%) for HPV16-driven OPC cases. The positive samples for each of
330 the targets contained a mean of 3,482 copies of HPV16 E6 (range 21.2 – 22,780 copies/ml),
331 3,431 copies of HPV16 E7 (range 10.6 – 28,880 copies/ml) and 42.7 copies of HPV33 E6
332 (range 10.6 – 74.8 copies/ml). Six samples were positive for HPV16 E6 and E7 and two
333 additional samples were positive for HPV16 E7 only, indicating a better performance of the
334 HPV16 E7 assay. The quantification of HPV16 E6 and E7 was highly correlated (Pearson's r >
335 0.99, fig. 3.3A). Cohen's kappa was 0.476 for HPV16 E6, 0.645 for HPV16 E7 and 1.0 for
336 HPV33 E6, indicating moderate to perfect agreement with tumor HPV status. There was no
337 correlation of HPV16 E7 copy numbers with BG copy numbers (Pearson's r = 0.01, fig. 3.3B).
338 There was a weak correlation between the number of BG copies detected and the plasma
339 volume (Pearson's r = 0.221, fig. 3.3C).

340 Both HPV16 assays had a specificity of 92% due to one false positive result in a sample from
341 an HPV33-driven OPC patient. One plasma sample from an HPV33-driven OPC case
342 reproducibly tested positive for HPV16 E6 (2/2 replicates on QIAcuity, 2/2 replicates on
343 QX200) and HPV16 E7 (2/2 replicates on QIAcuity) in addition to HPV33 E6 (2/2 replicates),
344 likely indicating that an HPV16 infection was missed in the tumor tissue.

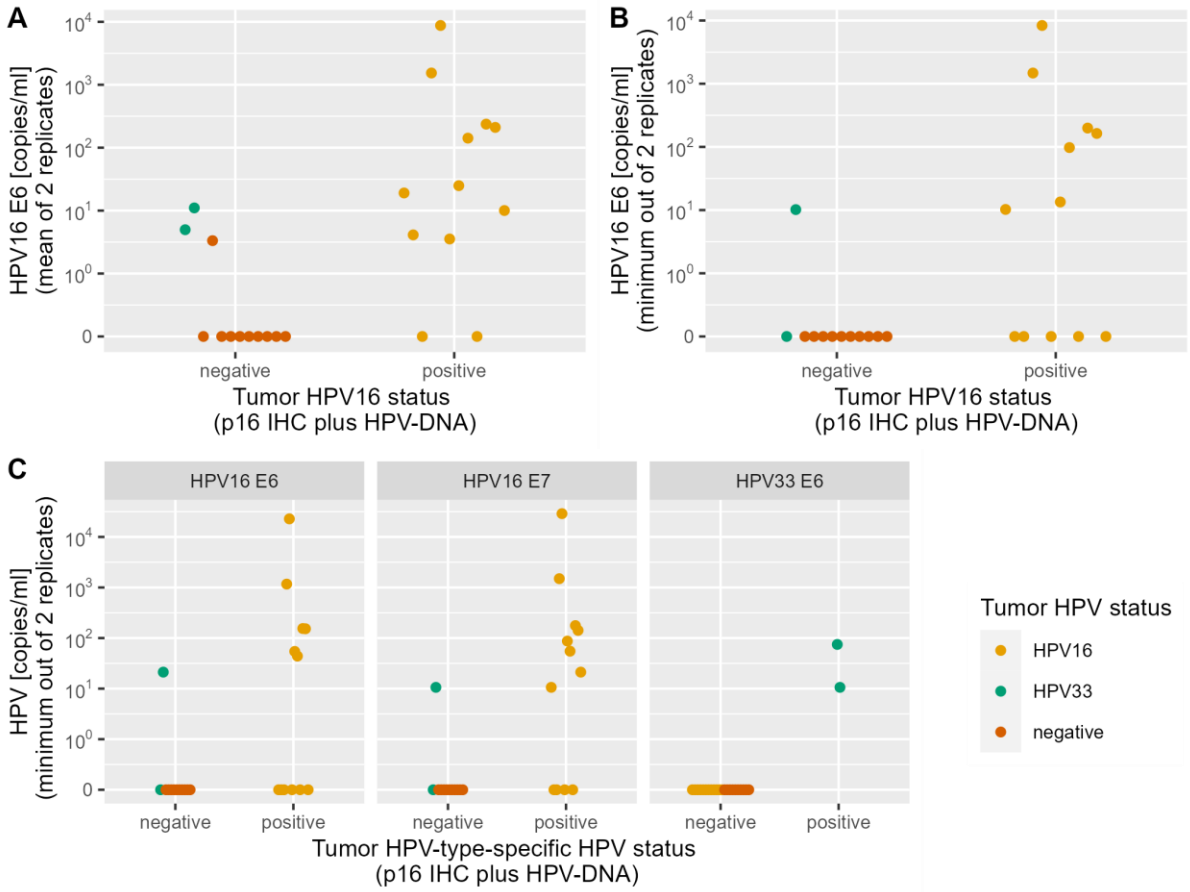


Figure 3.2: dPCR results of 24 patients from the training set (10 HPV-negative, 12 HPV16+, 2 HPV33+). Samples categorized as positive or negative depending on the HPV-type targeted by each experiment. (A) Copy number of HPV16 E6 cfDNA by QX200 ddPCR, showing the mean value of the replicates. (B) Copy number of HPV16 E6 cfDNA by QX200 ddPCR, showing the minimum value of the replicates. (C) Copy number of HPV16 E6, HPV16 E7 and HPV33 E6 cfDNA by QIAcuity dPCR, showing the minimum value of the replicates.

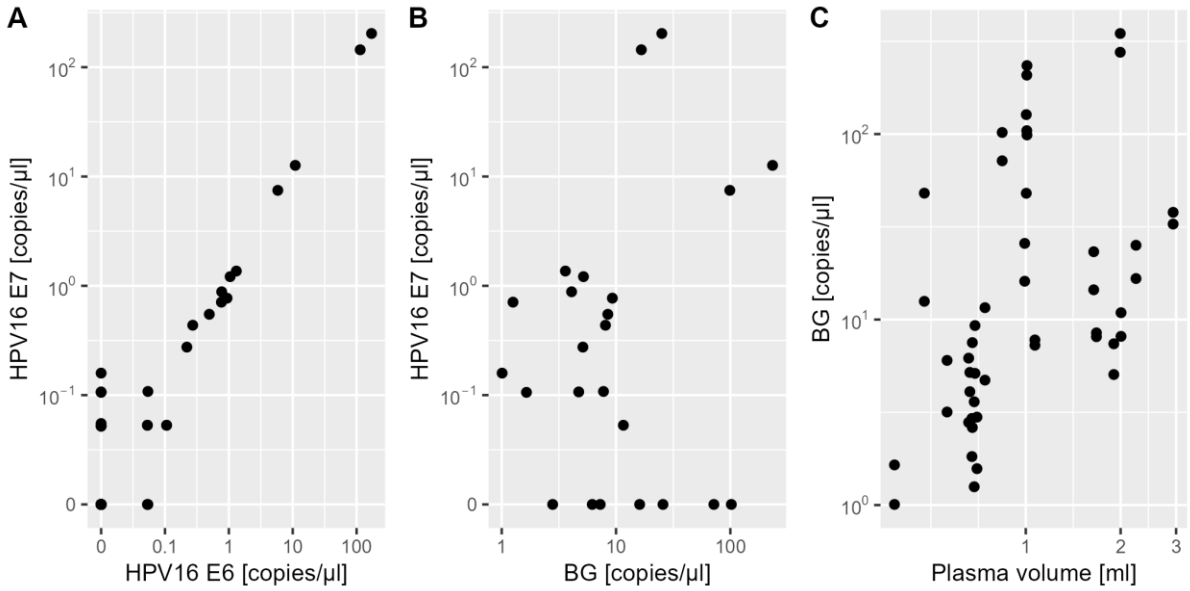
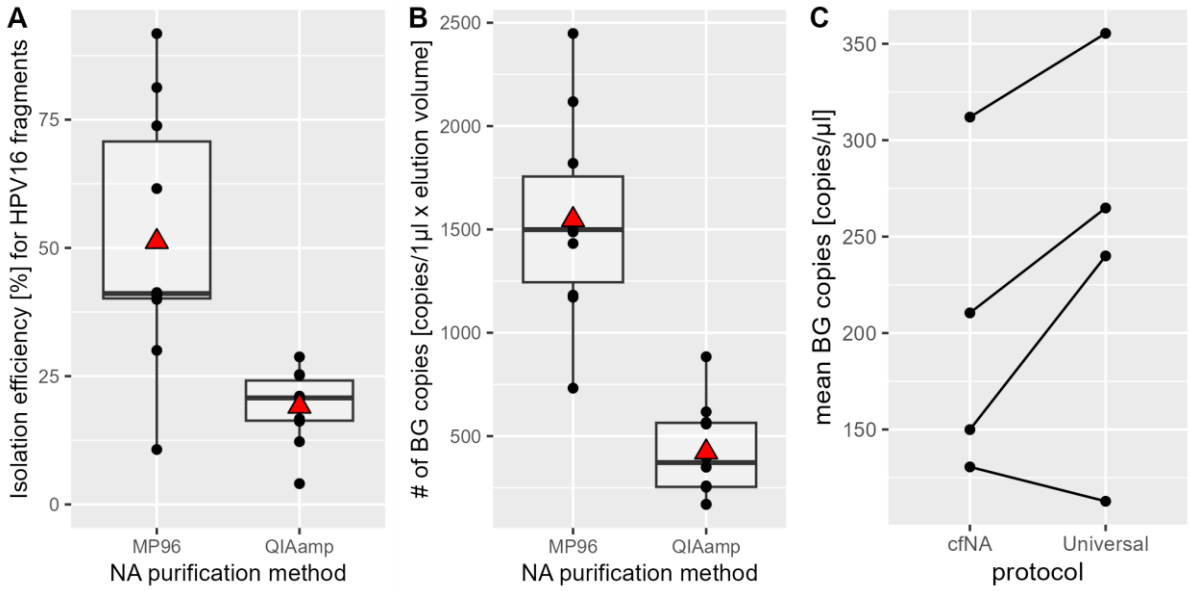


Figure 3.3: Number of copies of different target DNA detected in plasma samples from 24 OPC patients by QIAcuity dPCR. Each point represents one of two replicates from a sample. (A) Comparison of HPV16 E6 and HPV16 E7 copies detected in each sample. (B) Comparison of copy number of HPV16 E7 and BG copies in each sample. (C) Comparison of number of copies of BG detected and the input blood plasma volume.

3.2. **Optimization of cfDNA purification**

In order to optimally detect low concentrations of cfDNA in small volumes of less than 1 ml of blood plasma, the performance of different nucleic acid purification methods was analyzed. The automated *cfNA ss 2000* protocol for the MagNA Pure 96 was compared with the manual *QIAamp ccfDNA/RNA* kit using plasma from healthy volunteers with spiked-in HPV16 plasmid DNA. To ensure comparability, one plasma pool was split into multiple aliquots and tested on both platforms. Based on the quantification of spiked-in HPV16 DNA by QIAcuity dPCR, mean isolation efficiency was higher using the cfNA-protocol for the MagNA Pure 96 *cfNA ss 2000* (51%) compared to *QIAamp ccfDNA/RNA* kit (23%) (fig. 3.4A). Similarly, a higher number of BG copies from the native cfDNA were detected in the MagNA Pure 96 *cfNA ss 2000* extract (median 1,499 copies/ml) compared to the *QIAamp ccfDNA/RNA* extract (median 372 copies) (fig. 3.4B). When using the MagNA Pure 96 *cfNA ss 2000* protocol for isolating HPV DNA spiked into healthy donor EDTA plasma, as few as 10 copies of HPV DNA were detected in 8 of 8 replicates (100% sensitivity) in the subsequent QIAcuity dPCR, while with 5 copies the sensitivity was 75% (6/8 replicates). When cfDNA was isolated from sodium citrate plasma samples from healthy volunteers, the MagNA Pure 96 *cfNA ss 2000* protocol yielded fewer copies of BG per ml of plasma in 3 out of 4 samples than the *Viral NA Universal 500* protocol, despite using the same input volume in both protocols (fig. 3.4C). Consequently, the latter protocol was used for the sodium citrate plasma samples in the validation experiment.



378

379 **Figure 3.4: Performance of nucleic acid isolation from blood plasma. (A) Recovery rate of a**
380 **104 bp HPV16 fragment from EDTA plasma by MagNA Pure 96 *cfNA ss 2000 protocol* (MP96)**
381 **or *QIAamp ccfDNA/RNA kit* (QIAamp). The red triangle indicates the position of the group**
382 **mean. (B) Number of copies of native BG cfDNA isolated from aliquots of one EDTA plasma**
383 **pool by MP96 and QIAamp. (C) Comparison of MagNA Pure 96 *cfNA ss 2000 protocol* and the**
384 ***Viral NA Universal 500 protocol* for sodium citrate plasma.**

385 ## 3.3. **Effect of sample volume on quantification**

386 To gain insight about the effect of the sample volume on sensitivity, aliquots of 500 and
387 1000 µl were analyzed from 30 plasma samples from OPC patients (15 HPV16+, 2 HPV58+,
388 13 HPV-negative). The samples were analyzed in duplicates by QIAcuity dPCR.

389 All samples in both of the two groups (500 µl vs. 1000 µl) had sufficient human BG DNA.
390 When 1000 µl of plasma were used for cfDNA isolation, the HPV16 E7 dPCR had a sensitivity
391 of 80% for HPV16-driven tumors (n=12/15), while HPV16 E6 had a sensitivity of 73%
392 (n=11/15) (fig. 3.5). When 500 µl of sample was analyzed, the assay had a sensitivity of 73%
393 (n=11/15) for both HPV16 E6 and E7. In total only one HPV-driven OPC case was not
394 detected due to decreased input plasma volume. There were two false positive samples,
395 leading to a specificity of the assay of 87% (n=13/15). In one HPV58+ OPC plasma sample
396 very low amounts (2.5 copies/ml) of HPV16 E7-cfDNA only were detected in 500 µl extract
397 only. The other false positive plasma sample was from an OPC case where the
398 corresponding tumor tissue was HPV DNA-negative but p16-positive. This sample tested
399 concordantly positive for both HPV16 E6 and E7 cfDNA in both plasma aliquots of 500 and
400 1000 µl, likely indicating a missed HPV16 infection in the tumor tissue. Overall, the
401 quantification of cfDNA copies per ml was very consistent between the two sample volumes
402 for HPV16 E6 and E7 (r>=0.99), and was comparable for human BG (r=0.86).

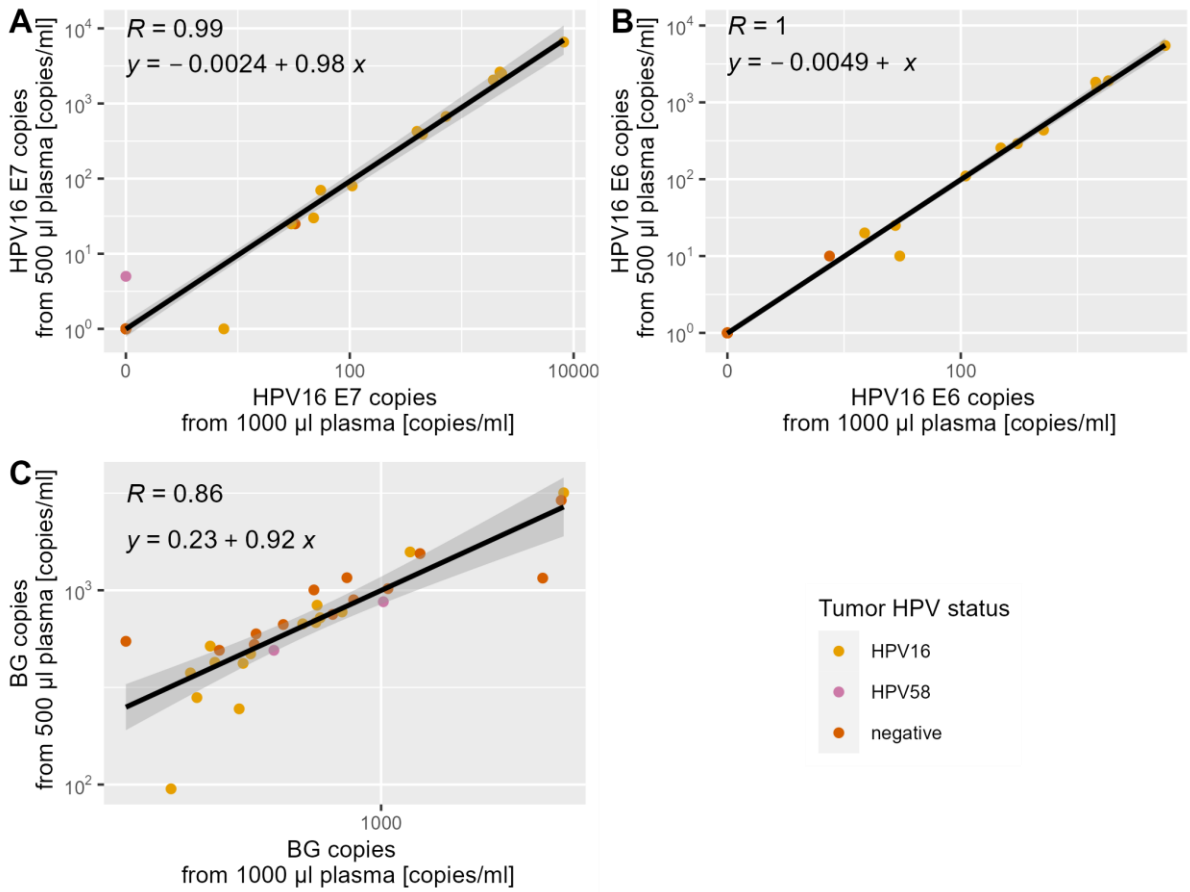


Figure 3.5: QIAcuity dPCR result of 30 OPC patient plasma samples (13 negative, 2 HPV58+, 15 HPV16+), comparing the concentration of each target (in copies/ml plasma) between cfDNA isolated from 500 µl (y-axis) and 1000 µl (x-axis) of plasma. (A) HPV16 E7, (B) HPV16 E6, (C) BG.

3.4. **Validation of final protocol in plasma samples from OPC patients**

For the final assessment of clinical sensitivity and specificity, a total of 65 plasma samples from OPC patients (34 HPV16+, 1 HPV33+, 2 HPV58+, 28 HPV negative) were analyzed, using the results from 1000 µl plasma volume where available. In the following analysis, HPV58-driven OPC cases were considered HPV-negative as the assay was not targeted to detect this HPV type. cfDNA was purified from plasma samples by MagNA Pure 96 *Viral NA Universal* protocol and applied in duplicates to QIAcuity dPCR. Only samples that tested reproducibly positive for a given target were considered positive for this target, based on results from the training set. Sample volumes analyzed for each sample ranged from 370 to 1000 µl (mean of 824 µl). The samples contained a median of 684 copies/ml of BG (range 223 – 80,982 copies/ml).

420 Samples diagnosed as HPV16 or HPV33-driven by pathology were classified correctly with a
421 sensitivity of 80% (28/35, 95% CI: 63 – 92%, tbl. 3.1) and with a specificity of 96.7% (29/30,
422 95% CI: 83 – 100%). The test had a Cohen's kappa of 0.756, indicating substantial agreement
423 with the tumor HPV status. In samples with 350 – 500 µl of plasma available, the sensitivity
424 of the assay was lower with 73.3% (11/15, 95% CI: 45 – 92%), while in samples with 750 –
425 1000 µl of plasma the sensitivity was 85% (17/20, 95% CI: 62 - 97%). A moderate correlation
426 between N-stage and HPV16 E7 copy numbers detected in plasma samples from HPV16-
427 driven OPCs was observed (r= 0.289), while there was a weak correlation with T-stage (r=
428 0.128). There was no correlation between BG copy numbers detected and the sensitivity of
429 the assay, but 5 out of 5 cases with more than 10,000 copies of beta-globin were HPV-
430 driven. The age of the sample had no effect on the number of copies of BG detected in the
431 plasma samples (r= -0.076).

432 **Table 3.1: Performance of QIAcuity dPCR of plasma cfDNA compared to tumor HPV status.**

cfDNA[1]	***Tumor HPV status (HPV DNA + p16-IHC)***	
	HPV16 +	**HPV16 -**
HPV16 E7+	27	2
HPV16 E7-	7	29
	HPV16 +	**HPV16 -**
HPV16 E6+	22	2
HPV16 E6-	12	29
	HPV33 +	**HPV33 -**
HPV33 E6+	1	0
HPV33 E6-	0	64
cfDNA[1]	**HPV16/33 +**	**HPV16/33 -**
HPV+[2]	28	1
HPV-[2]	7	29
Sensitivity	0.800 (95% CI: 0.63 - 0.92)	
Specificity	0.967 (95% CI: 0.83 - 1.00)	
Kappa	0.756 (95% CI: 0.68 - 0.84)	

433 [1] HPV detected in plasma by QIAcuity dPCR, [2] HPV+/- summarizes the results of HPV16 E7,
434 HPV16 E6 and HPV33 E6 dPCR; the overall result was positive if any of the three targets was
435 detected in both duplicates for a sample.

436 4. Discussion

437 Early detection of OPC holds the potential to improve patient's quality of life due to early
438 therapeutic intervention requiring less intensive therapy. For this purpose, novel diagnostic
439 molecular biomarkers are needed using minimally invasive protocols. HPV cfDNA is a
440 promising candidate that is successfully being used in therapy surveillance of HPV-driven
441 OPC already. To evaluate this biomarker for early diagnosis a highly sensitive and specific
442 assay is needed that can also be successfully applied to archived material, where usually
443 only small amounts of plasma are available. Here, we describe a digital PCR assay targeting
444 HPV16 and HPV33.

445 To optimize the assay at every step from DNA isolation from plasma to DNA detection and
446 quantification, multiple comparisons of methods were conducted. Due to the superior
447 cfDNA isolation efficiency of the MagNA Pure 96 over the QIAamp kit, the MagNA Pure 96
448 was used in the final validation set of samples. While the QX200 and QIAcuity showed
449 comparable analytical and clinical sensitivity, the QIAcuity offers higher multiplexing
450 capacity, which enables the targeting of multiple HPV types at once. In addition, the effect
451 of the sample volume on the sensitivity was evaluated. Our results suggest that it is feasible
452 to use lower input volumes than seen in most studies, as using only up to 1 ml of plasma in
453 the validation experiment yielded a sensitivity of 80%, which is comparable to other studies
454 which achieved 70 – 96% sensitivity working with 2 – 4 ml of plasma per sample [13] [15]
455 [18] [17]. Since even in samples with 500 µl or less plasma available the sensitivity was still
456 73.3%, the assay is suitable for use in archived samples from large collections in biobanks,
457 where usually only small volumes are available.

458 The increase in sensitivity from the training set (71%) to the validation set (80%) could be
459 explained by an increased PCR input volume from 10 to 20 µl, due to increased sample
460 volume requirements in the training set for the comparison of two different dPCR methods,
461 and the use of an optimized cfDNA purification method in the validation set. The
462 unexpectedly low cfDNA yield in samples with high input volumes prompted a more in-
463 depth comparison of different cfDNA isolation protocols for these samples. This led to the
464 conclusion that the *Viral NA Universal* protocols give a higher yield in these samples than
465 the cfDNA protocols, so that the protocol used for cfDNA isolation was changed in the
466 validation set.

467 The specificity of the QIAcuity dPCR in the final validation set was 96.6%. The one
468 reproducibly false HPV16 E6 and E7 cfDNA positive plasma sample by dPCR was from a
469 patient with a tumor that was tested HPV DNA-negative and p16-positive by a pathologist.
470 The same plasma sample has been previously tested positive for antibodies against multiple
471 HPV16 antigens using multiplex serology, which is an assay that has a very high specificity
472 [34] [35]. Therefore it seems likely that the HPV DNA in the tumor tissue has been missed by
473 the reference assay, and it could be reasonably assumed that the tumor was actually HPV-
474 driven, which would set the actual specificity of the assay to 100%.

475 In the training set and the experiment for the direct comparison of cfDNA isolated from
476 different volumes, three additional false positive cases emerged. In contrast to the HPV DNA
477 result of corresponding tumor tissue samples, additional HPV16 DNA was detected in three
478 plasma samples from HPV33 or HPV58-driven OPC patients. This could be explained by
479 either lack of type-specificity of the assay when applied to cfDNA, which seems unlikely as
480 there was no cross-reactivity in cervical swab samples positive for any of these HPV-types,
481 or previously undetected co-infections driving the cancer in these cases due to sensitivity
482 limitations of the reference test. In cervical swabs within this study there were multiple
483 cases of co-infection with these HPV types, which have also been frequently observed in
484 other studies [36], although whether these findings are applicable to the oropharynx
485 requires further investigation.

486 Although HPV16 is the most common HPV type among HPV-driven OPC, one limitation of
487 the assay is the small number of HPV types covered. In our study, HPV58 was present almost
488 as frequently as HPV33. However, only HPV33 was included in the assay due to limited
489 multiplexing spectrum of the dPCR. Since all HPV16 E6 positive sample were also HPV16 E7
490 positive, in future studies HPV16 E6 could be replaced by primers and probes detecting
491 other HPV types.

492 Since therapy surveillance is a well-established use for HPV cfDNA detection, a more
493 interesting question to answer would be the kinetics of HPV cfDNA before diagnosis of HPV-
494 driven cancer. The possibility of using lower input volumes may open up new sample
495 sources that would not be able to provide the volumes usually used for cfDNA isolation. The
496 assay presented in this study is sensitive and specific even for these smaller input volumes
497 and thus could help exploring the utility of HPV cfDNA in early detection of HPV-driven
498 cancers, which could be feasible due to only requiring minimally invasive sampling.

499 # 5. Conclusions

500 Briefly, we established a novel dPCR detecting HPV16 E6 and E7, HPV33 E6 and the internal
501 BG control on the QIAcuity in archived, low volume plasma samples after using automated
502 DNA extraction with the MagNA Pure 96. The dPCR was shown to detect sensitively and
503 specifically HPV16 and HPV33. Follow-up studies in larger cohorts are required to explore
504 the use of HPV cfDNA in early detection of HPV-driven OPC.

505 # 6. Funding sources

506 This study was funded by intramural DKFZ funding.

507 # References

508 [1] Szymonowicz KA, Chen J. Biological and clinical aspects of HPV-related cancers.
509 Cancer Biol Med 2020;17:864–78. https://doi.org/10.20892/j.issn.2095-
510 3941.2020.0370.

511 [2] Pytynia KB, Dahlstrom KR, Sturgis EM. Epidemiology of HPV-associated
512 oropharyngeal cancer. Oral Oncol 2014;50:380–6.
513 https://doi.org/10.1016/j.oraloncology.2013.12.019.

514 [3] Hong AM, Martin A, Armstrong BK, Lee CS, Jones D, Chatfield MD, et al. Human
515 papillomavirus modifies the prognostic significance of t stage and possibly n stage in
516 tonsillar cancer. Ann Oncol 2013;24:215–9.
517 https://doi.org/10.1093/annonc/mds205.

518 [4] Mahajan A. Practical issues in the application of p16 immunohistochemistry in
519 diagnostic pathology. Hum Pathol 2016;51:64–74.
520 https://doi.org/10.1016/j.humpath.2015.12.021.

521 [5] You EL, Henry M, Zeitouni AG. Human papillomavirus-associated oropharyngeal
522 cancer: Review of current evidence and management. Curr Oncol 2019;26:119–23.
523 https://doi.org/10.3747/co.26.4819.

524 [6] Pinkiewicz M, Dorobisz K, Zatonski T. Human papillomavirus-associated head and
525 neck cancers. Where are we now? A systematic review. Cancer Manag Res
526 2022;14:3313–24. https://doi.org/10.2147/CMAR.S379173.

527 [7] Hirth JM, Chang M, Resto VA, Group HPVS. Prevalence of oral human papillomavirus
528 by vaccination status among young adults (18-30years old). Vaccine 2017;35:3446–
529 51. https://doi.org/10.1016/j.vaccine.2017.05.025.

530 [8] Damgacioglu H, Sonawane K, Chhatwal J, Lairson DR, Clifford GM, Giuliano AR, et al.
531 Long-term impact of HPV vaccination and COVID-19 pandemic on oropharyngeal
532 cancer incidence and burden among men in the USA: A modeling study. Lancet Reg
533 Health Am 2022;8:100143. https://doi.org/10.1016/j.lana.2021.100143.

534 [9] Jahr S, Hentze H, Englisch S, Hardt D, Fackelmayer FO, Hesch RD, et al. DNA
535 fragments in the blood plasma of cancer patients: Quantitations and evidence for
536 their origin from apoptotic and necrotic cells. Cancer Res 2001;61:1659–65.

537 [10] Lui YY, Chik KW, Chiu RW, Ho CY, Lam CW, Lo YM. Predominant hematopoietic origin
538 of cell-free DNA in plasma and serum after sex-mismatched bone marrow
539 transplantation. Clin Chem 2002;48:421–7.

540 [11] Nguyen B, Meehan K, Pereira MR, Mirzai B, Lim SH, Leslie C, et al. A comparative
541 study of extracellular vesicle-associated and cell-free DNA and RNA for HPV
542 detection in oropharyngeal squamous cell carcinoma. Sci Rep 2020;10:6083.
543 https://doi.org/10.1038/s41598-020-63180-8.

544 [12] Cao Y, Haring CT, Brummel C, Bhambhani C, Aryal M, Lee C, et al. Early HPV ctDNA
545 kinetics and imaging biomarkers predict therapeutic response in p16+ oropharyngeal
546 squamous cell carcinoma. Clin Cancer Res 2022;28:350–9.
547 https://doi.org/10.1158/1078-0432.CCR-21-2338.

548 [13] Chera BS, Kumar S, Beaty BT, Marron D, Jefferys S, Green R, et al. Rapid clearance
549 profile of plasma circulating tumor HPV type 16 DNA during chemoradiotherapy
550 correlates with disease control in HPV-associated oropharyngeal cancer. Clin Cancer
551 Res 2019;25:4682–90. https://doi.org/10.1158/1078-0432.CCR-19-0211.

552 [14] Siravegna G, O'Boyle CJ, Varmeh S, Queenan N, Michel A, Stein J, et al. Cell-free HPV
553 DNA provides an accurate and rapid diagnosis of HPV-associated head and neck
554 cancer. Clin Cancer Res 2022;28:719–27. https://doi.org/10.1158/1078-0432.CCR-
555 21-3151.

556 [15] Haring CT, Bhambhani C, Brummel C, Jewell B, Bellile E, Heft Neal ME, et al. Human
557 papilloma virus circulating tumor DNA assay predicts treatment response in
558 recurrent/metastatic head and neck squamous cell carcinoma. Oncotarget
559 2021;12:1214–29. https://doi.org/10.18632/oncotarget.27992.

560 [16] Tewari SR, D'Souza G, Troy T, Wright H, Struijk L, Waterboer T, et al. Association of
561 plasma circulating tumor HPV DNA with HPV-related oropharynx cancer. JAMA

562 Otolaryngol Head Neck Surg 2022;148:488–9.
563 https://doi.org/10.1001/jamaoto.2022.0159.

564 [17] Mattox AK, D'Souza G, Khan Z, Allen H, Henson S, Seiwert TY, et al. Comparison of
565 next generation sequencing, droplet digital PCR, and quantitative real-time PCR for
566 the earlier detection and quantification of HPV in HPV-positive oropharyngeal
567 cancer. Oral Oncol 2022;128:105805.
568 https://doi.org/10.1016/j.oraloncology.2022.105805.

569 [18] Damerla RR, Lee NY, You D, Soni R, Shah R, Reyngold M, et al. Detection of early
570 human papillomavirus-associated cancers by liquid biopsy. JCO Precis Oncol 2019;3.
571 https://doi.org/10.1200/PO.18.00276.

572 [19] Holzinger D, Wichmann G, Baboci L, Michel A, Hofler D, Wiesenfarth M, et al.
573 Sensitivity and specificity of antibodies against HPV16 E6 and other early proteins for
574 the detection of HPV16-driven oropharyngeal squamous cell carcinoma. Int J Cancer
575 2017;140:2748–57. https://doi.org/10.1002/ijc.30697.

576 [20] Kreimer AR, Ferreiro-Iglesias A, Nygard M, Bender N, Schroeder L, Hildesheim A, et
577 al. Timing of HPV16-E6 antibody seroconversion before OPSCC: Findings from the
578 HPVC3 consortium. Ann Oncol 2019;30:1335–43.
579 https://doi.org/10.1093/annonc/mdz138.

580 [21] Waterboer T, Brenner N, Gallagher R, Hillman RJ, Jin F, Grulich A, et al. Early
581 detection of human papillomavirus-driven oropharyngeal cancer using serology from
582 the study of prevention of anal cancer. JAMA Oncol 2020;6:1806–8.
583 https://doi.org/10.1001/jamaoncol.2020.4527.

584 [22] Busch CJ, Hoffmann AS, Viarisio D, Becker BT, Rieckmann T, Betz C, et al. Detection of
585 stage i HPV-driven oropharyngeal cancer in asymptomatic individuals in the hamburg
586 city health study using HPV16 E6 serology - a proof- of-concept study.
587 Eclinicalmedicine 2022;53. https://doi.org/ARTN
588 10165910.1016/j.eclinm.2022.101659.

589 [23] Roda Husman AM de, Walboomers JM, Brule AJ van den, Meijer CJ, Snijders PJ. The
590 use of general primers GP5 and GP6 elongated at their 3' ends with adjacent highly
591 conserved sequences improves human papillomavirus detection by PCR. J Gen Virol
592 1995;76 (Pt 4):1057–62. https://doi.org/10.1099/0022-1317-76-4-1057.

593 [24] Hofler D, Bohmer G, Wasielewski R von, Neumann H, Halec G, Holzinger D, et al.
594 HPV16 RNA patterns defined by novel high-throughput RT-qPCR as triage marker in

595 HPV-based cervical cancer precursor screening. Gynecologic Oncology
596 2015;138:676–82. https://doi.org/10.1016/j.ygyno.2015.06.039.

597 [25] Schmitt M, Bravo IG, Snijders PJ, Gissmann L, Pawlita M, Waterboer T. Bead-based
598 multiplex genotyping of human papillomaviruses. J Clin Microbiol 2006;44:504–12.
599 https://doi.org/10.1128/JCM.44.2.504-512.2006.

600 [26] Baboci L, Holzinger D, Boscolo-Rizzo P, Tirelli G, Spinato R, Lupato V, et al. Low
601 prevalence of HPV-driven head and neck squamous cell carcinoma in north-east
602 italy. Papillomavirus Res 2016;2:133–40. https://doi.org/10.1016/j.pvr.2016.07.002.

603 [27] Allawi HT, SantaLucia Jr J. Thermodynamics of internal c.t mismatches in DNA.
604 Nucleic Acids Res 1998;26:2694–701. https://doi.org/10.1093/nar/26.11.2694.

605 [28] Beaudenon S, Kremsdorf D, Croissant O, Jablonska S, Wain-Hobson S, Orth G. A novel
606 type of human papillomavirus associated with genital neoplasias. Nature
607 1986;321:246–9. https://doi.org/10.1038/321246a0.

608 [29] Seedorf K, Krammer G, Durst M, Suhai S, Rowekamp WG. Human papillomavirus type
609 16 DNA sequence. Virology 1985;145:181–5. https://doi.org/10.1016/0042-
610 6822(85)90214-4.

611 [30] R-Core-Team. R: A language and environment for statistical computing 2022.

612 [31] Wickham H, Averick M, Bryan J, Chang W, D'Agostino McGowan L, Romain Françoi R,
613 et al. Welcome to the tidyverse. Journal of Open Source Software 2019;4:1686.
614 https://doi.org/10.21105/joss.01686.

615 [32] Stevenson M, Sergeant E. epiR: Tools for the analysis of epidemiological data 2023.

616 [33] Kuhn M. Building predictive models in r using the caret package. Journal of Statistical
617 Software 2008;28:1–26. https://doi.org/10.18637/jss.v028.i05.

618 [34] Hibbert J, Halec G, Baaken D, Waterboer T, Brenner N. Sensitivity and specificity of
619 human papillomavirus (HPV) 16 early antigen serology for HPV-driven oropharyngeal
620 cancer: A systematic literature review and meta-analysis. Cancers (Basel) 2021;13.
621 https://doi.org/10.3390/cancers13123010.

622 [35] Laban S, Gangkofner DS, Holzinger D, Schroeder L, Eichmuller SB, Zornig I, et al.
623 Antibody responses to cancer antigens identify patients with a poor prognosis
624 among HPV-positive and HPV-negative head and neck squamous cell carcinoma
625 patients. Clin Cancer Res 2019;25:7405–12. https://doi.org/10.1158/1078-0432.CCR-
626 19-1490.

[36] Kim M, Park NJ, Jeong JY, Park JY. Multiple human papilloma virus (HPV) infections are associated with HSIL and persistent HPV infection status in korean patients. Viruses 2021;13. https://doi.org/10.3390/v13071342.

Appendix

To be written