



## OPEN ACCESS

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/ijgc-2022-003911>).

For numbered affiliations see end of article.

## Correspondence to

Majke H.D. van Bommel, Obstetrics & Gynaecology, Radboudumc, Nijmegen, The Netherlands; majke.vanbommel@radboudumc.nl

Received 9 August 2022

Accepted 17 October 2022

Published Online First

16 November 2022



© IGCS and ESGO 2022. Re-use permitted under CC BY. Published by BMJ.

**To cite:** van Bommel MH.D, Pijnenborg JM.A, van der Putten LJM, *et al.* *Int J Gynecol Cancer* 2022;**32**:1568–1575.

# Diagnostic accuracy of mutational analysis along the Müllerian tract to detect ovarian cancer

Majke H.D. van Bommel <sup>1</sup>, Johanna M.A. Pijnenborg, <sup>1</sup> Louis J M van der Putten, <sup>1</sup> Johan Bulten, <sup>2</sup> Marc P.L.M. Snijders, <sup>3</sup> Heidi V.N. Küsters-Vandeveldt, <sup>4</sup> Sanne Sweegers, <sup>2</sup> M. Caroline Vos, <sup>5</sup> Marjolein J.L. Ligtenberg, <sup>2,6</sup> Astrid Eijkelenboom, <sup>2</sup> Joanne A de Hullu, <sup>1</sup> Casper Reijnen <sup>7</sup>

## ABSTRACT

**Objective** Ovarian cancer is known for its poor prognosis, which is mainly due to the lack of early symptoms and adequate screening options. In this study we evaluated whether mutational analysis in cervicovaginal and endometrial samples could assist in the detection of ovarian cancer.

**Methods** In this prospective multicenter study, we included patients surgically treated for either (suspicion of) ovarian cancer or for a benign gynecological condition (control group). A cervicovaginal self-sample, a Papanicolaou (Pap) smear, a pipelle endometrial biopsy, and the surgical specimen were analyzed for (potentially) pathogenic variants in eight genes (*ARID1A*, *CTNNB1*, *KRAS*, *MTOR*, *PIK3CA*, *POLE*, *PTEN*, and *TP53*) using single-molecule molecular inversion probes. Sensitivity and specificity were calculated to assess diagnostic accuracy.

**Results** Based on surgical histology, our dataset comprised 29 patients with ovarian cancer and 32 controls. In 83% of the patients with ovarian cancer, somatic (potentially) pathogenic variants could be detected in the final surgical specimen, of which 71% included at least a *TP53* variant. In 52% of the ovarian cancer patients, such variants could be detected in either the self-sample, Pap smear, or pipelle. The Pap smear yielded the highest diagnostic accuracy with 26% sensitivity (95% CI 10% to 48%). Overall diagnostic accuracy was low and was not improved when including *TP53* variants only.

**Conclusions** Mutational analysis in cervicovaginal and endometrial samples has limited accuracy in the detection of ovarian cancer. Future research with cytologic samples analyzed on methylation status or the vaginal microbiome may be relevant.

## INTRODUCTION

Epithelial ovarian cancer is the most lethal gynecological cancer.<sup>1</sup> Patients generally present with advanced-stage disease leading to a 5-year survival of approximately 45%,<sup>2</sup> mainly due to the absence of early symptoms and reliable screening methods.<sup>3</sup> By contrast, survival for the limited number of patients with localized disease is around 92%, suggesting that early detection of epithelial ovarian cancer could substantially improve prognosis.<sup>2</sup>

Epithelial ovarian cancer is thought to develop from tissues embryologically derived from the Müllerian

## WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ DNA variants related to ovarian cancer can be detected along the Müllerian tract. In this study we compared DNA pathogenic variants in cervicovaginal self-samples, Papanicolaou (Pap) smears, and pipelle endometrial biopsies with the pathogenic variants in the surgical specimen in patients with ovarian cancer and control patients. We assessed diagnostic accuracy of detecting ovarian cancer with those samples.

## WHAT THIS STUDY ADDS

⇒ We found that diagnostic accuracy was low for cervicovaginal self-samples, Pap smears, and endometrial biopsies. Thus, the samples assessed in this way cannot be used for early detection of ovarian cancer.

## HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE, OR POLICY

⇒ This study contributes to further unraveling the oncogenesis of ovarian carcinoma and assists in research regarding the urgently needed detection of ovarian cancer.

ducts (fallopian tubes, uterus, upper part of the vagina) with the ovaries secondarily involved. Nowadays, ovarian, fallopian tubal, and/or the peritoneal malignancies are considered collectively as ovarian carcinomas, of which approximately 75% are high-grade serous carcinomas. There is compelling evidence that high-grade serous carcinoma originates in the fallopian tubes,<sup>4</sup> potentially offering new strategies for ovarian cancer prevention and early detection.

Screening for ovarian cancer using transvaginal sonography and cancer antigen 125 (CA125) has been proven ineffective in the general population<sup>3</sup> and in women at increased inherited risk.<sup>5</sup> Lately, instead of focusing on macroscopic changes, there is increasing interest in detecting microscopic (pre)malignant cells that detach along the Müllerian ducts. Interest in DNA analysis in cytological samples is growing since (cell-free) DNA variants can be detected in cytological samples even without the presence of tumor cells. Kinde *et al* extracted DNA from a Papanicolaou (Pap)

smear and found a sensitivity of detecting ovarian cancer of 41%, particularly driven by mutated Tumor Protein 53 (*TP53*) variants.<sup>6</sup> When combining mutational analysis with DNA extracted from a Pap smear with plasma, sensitivity improved to 63%. An intra-uterine brush to sample DNA closer to the primary source increased sensitivity.<sup>7</sup> Currently, uterine and tubal lavage to detect early-stage ovarian cancer is being investigated (NCT 02039388). The first results are promising as ovarian cancer cells could be collected in 24 of 30 patients with ovarian cancer, and mainly *TP53* mutations could be identified.<sup>8</sup>

The above findings support the presence of ovarian cancer cells along the Müllerian tract, which could potentially be detected with minimally invasive sampling methods. Therefore, we investigated the diagnostic accuracy of detecting ovarian cancer by assessing DNA pathogenic variants in cervicovaginal and endometrial samples and comparing them with the pathogenic variants found in the tumor itself.

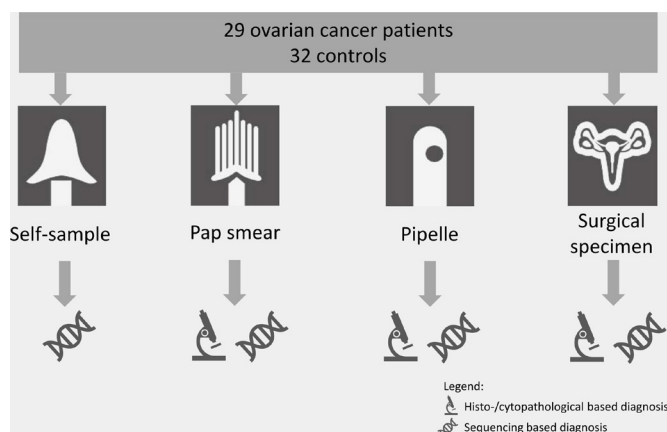
## METHODS

### Design and Population

This prospective observational multicenter study included consecutive patients undergoing surgery for high suspicion of ovarian cancer or for a benign gynecological condition (control group) in three Dutch hospitals: Radboud University Medical Center, Nijmegen; Canisius-Wilhelmina Hospital, Nijmegen; and Elisabeth-TweeSteden Hospital, Tilburg. Suspicion was based on a Risk Malignancy Index >200, the presence of ascites, peritoneal depositions, omental cake, or laparoscopic evaluation. Inclusion criteria were adult age and surgery between December 2013 and January 2017 in a participating hospital. Exclusion criteria were a history of pelvic radiotherapy or previous hysterectomy. Ethical approval was obtained in all hospitals (Study Number 2013/451) and each patient signed informed consent. The study was prospectively registered at the Dutch Trial Registry (NTR4299) and performed according to the STARD guidelines for Standards for the Reporting of Diagnostic accuracy studies. Patients with endometrial cancer were included as well. Their results have been published previously.<sup>9</sup>

### Data Collection

Four specimens were collected from each patient: cervicovaginal self-sample, Pap smear, pipelle endometrial biopsy, and surgical



**Figure 1** Study flowchart.

sample (ovarian tissue) (Figure 1). Samples were collected in the aforementioned order by the operating gynecologist on the day of surgery. Demographic information was extracted from medical records.

### Pathogenic Variant Analysis

The complete workflow is provided in Online supplemental document 1. Briefly, DNA was extracted from the four specimens and analyzed using single-molecule molecular inversion probes-based sequencing on a NextSeq500 device (Illumina, San Diego, California, USA), as previously described.<sup>10</sup> The single-molecule molecular inversion probes were constructed to highlight hotspots in the oncogenes relevant in ovarian and endometrial cancer: Catenin Beta 1 (*CTNNB1*), Kirsten rat sarcoma virus (*KRAS*), mammalian target of rapamycin (*MTOR*), Phosphatidylinositol-4,5-Bisphosphate 3-Kinase Catalytic Subunit Alpha (*PIK3CA*), and Polymerase  $\epsilon$  (*POLE*); and all coding and splice site sequences of the tumor suppressor genes: AT-rich interactive domain-containing protein 1A (*ARID1A*), Phosphatase and tensin homolog (*PTEN*), and Tumor Protein 53 (*TP53*). Genes were chosen based on the genetic characteristics of ovarian and endometrial cancer as described in The Cancer Genome Atlas (TCGA).<sup>11–13</sup> Variants were categorized according to the following classes: 1, benign; 2, likely benign; 3, variant of unknown significance; 4, likely pathogenic, and 5, pathogenic. The last three classes were considered (potentially) pathogenic.

All surgical ovarian samples and pipelle endometrial biopsies were analyzed for the presence of the (potentially) pathogenic variants with a variant allele frequency of  $\geq 3\%$  and a minimal number of five variant reads (equal to three unique genomic DNA molecules). Additionally, the pipelle data were evaluated at the positions with known pathogenic variants in the surgical specimen with a cut-off of five unique variant reads and no minimal variant allele frequency. For evaluation of the Pap smears and self-samples, two independent library preparations were analyzed using a minimal variant allele frequency of 1% as we expected low variant allele frequencies with the samples mainly containing healthy endocervical cells and few tumor cells.

### Data Analysis

Baseline data were analyzed descriptively and differences between groups were analyzed using a t-test, Mann–Whitney U test, or  $\chi^2$  test. To measure diagnostic accuracy, we calculated sensitivity and specificity. First, we calculated the detection rates of (potentially) pathogenic variants per sample among all patients. Second, we measured diagnostic accuracy among all patients using the detection of (potentially) pathogenic variants in their surgical sample as gold standard. Third, we focused on *TP53* pathogenic variants as these variants are primarily related to ovarian cancer.<sup>14</sup>

## RESULTS

Specimens were collected from 37 patients with ovarian cancer and 32 controls. Eight patients with ovarian cancer were excluded because sequencing of the surgical specimen (the gold standard) was unsuccessful: the coverage was too low to detect any potential variant. Thus, 29 patients with ovarian cancer and 32 controls were included. Table 1 shows the characteristics of the patients.

## Original research

**Table 1** Baseline characteristics

	Ovarian cancer patients (n=29)	Controls (n=32)	P value
Age, years	66 (32–83)	57 (45–82)	0.063
Body mass index, kg/m <sup>2</sup>	25 (19–32)	23 (20–44)	0.644
CA125, kIU/L	360 (8–2196)	20 (9–142)	<0.001
Menopausal status			0.321
Pre-menopausal	4 (14)	8 (25)	
Post-menopausal	25 (86)	23 (72)	
Unknown	0	1 (3)	
Histology			
High-grade serous	20 (69)		
Endometrioid	3 (10)		
Clear cell	1 (3)		
Malignant mixed Müllerian tumor	1 (3)		
Mixed*	4 (14)		
Myoma		9 (28)	
Cystadenoma ovarii		10 (31)	
Fibroma/teratoma ovarii		7 (22)	
Other†		6 (19)	
FIGO stage		NA	
IA	1 (3)		
IB	0		
IC	2 (7)		
IIA	2 (7)		
IIB	2 (7)		
IIC	0		
IIIA	1 (3)		
IIIB	4 (14)		
IIIC	14 (48)		
IV	3 (10)		

Values presented as median (range) or N (%).

\*Mixed histology included two clear cell/serous, one clear cell/endometrioid, and one serous/endometrioid.

†Other histology included one adenomyosis, one inflammation, two prolapse, and two normal.

CA125, cancer antigen 125; FIGO, International Federation of Gynecology and Obstetrics.

### Surgical Specimens

In the tumors of 29 patients with ovarian cancers, 79% of the exons had a mean coverage of >250 reads, reflecting a 95% probability of detecting a variant, 'adequately sequenced' (see Online supplemental table 2).<sup>10</sup> We detected at least one pathogenic variant in 24 patients (83%). In total, 34 (potentially) pathogenic variants were detected (Figure 2A). Among the 24 patients with ovarian cancer with pathogenic variants, 17 had a *TP53* variant (71%). Fifteen of these 17 had high-grade serous carcinoma, one had a clear cell/endometrioid ovarian cancer, and one a clear cell/serous ovarian cancer. Other detected variants were: *PIK3CA* (21%), *CTNNB1* (14%), *PTEN* (10%), *KRAS* (7%), *ARID1A* (3%), and *MTOR* (3%) (Table 2). Of the five patients without a detected variant in their surgical sample (17%), three had high-grade serous carcinoma, one had a clear cell carcinoma, and one had clear cell/serous histology. Among the controls, two patients (6%), both with a mucinous cystadenoma,

were found to have a *KRAS* variant in their surgical specimen. Individual characteristics are shown in Online supplemental table 3.

### Cervicovaginal Samples (Self-Samples and Pap Smears)

Analysis of the cervicovaginal self-samples of 27 patients with ovarian cancer identified six patients (22%) with a total of seven (potentially) pathogenic variants: *PIK3CA* (n=3), *TP53* (n=2), and *ARID1A* (n=2); 98% of the exons were 'adequately sequenced'. When evaluating overlapping variants (variants that were detected in more than one specimen), we found that two variants were found in all samples but the surgical specimen (one *TP53* variant, one *PIK3CA*); and one variant was found in the surgical specimen, self-sample and pipelle but not in the Pap smear (*ARID1A*). Four variants were solely found in the self-sample (see Figure 2B and Online supplemental table 4). One control had an *ARID1A* variant in her self-sample that was not detected in her other specimens.



**Figure 2** Overview of detected (potentially) pathogenic variants in the surgical specimens (A), self-samples (B), Pap smears (C), and pipelle endometrial biopsies (D) and visualization of overlapping and non-overlapping variants between the surgical specimen and self-samples (B), Pap smears (C), and pipelle (D) in patients with ovarian cancer and controls.









Results of the Pap smears were available for 26 patients with ovarian cancer; exons ‘adequately sequenced’: 98%. Among them, 10 patients (38%) had one or more (potentially) pathogenic variants. In total, we detected 15 variants of which eight were overlapping with the variants in the surgical specimen (53%) (see Figure 2C and Online supplemental table 4). The eight overlapping variants were detected in six patients. Four of these six patients had high-grade serous carcinoma (of which three had a *TP53* variant (two stage 3C and one stage 2A) and one had a *PIK3CA*, *KRAS* and *PTEN* variant (high-grade serous carcinoma stage 3C); one had an endometrioid carcinoma stage 2A (and *CTNMB1* variant); and one had a clear cell/endometrioid carcinoma stage 1C (and *PIK3CA*

variant)). Two control patients had both a *PIK3CA* variant solely in their Pap smear.

#### Pipelle Endometrial Biopsies

A pipelle was available for 23 patients with ovarian cancer; 71% of the exons were ‘adequately sequenced’. Analysis showed seven patients (30%) with a total of 10 pathogenic variants: four *PIK3CA*, two *TP53*, two *KRAS*, one *PTEN*, and one *ARID1A*. We found one overlapping variant between the surgical specimens and the pipelles (10%), which was an *ARID1A* variant in a woman with high-grade serous carcinoma stage 3B who had this variant in all specimens but the Pap smear (see Figure 2D and Online supplemental

**Table 2** Number of (potentially) pathogenic variants in the various specimens

<b>Ovarian cancer patients (n=29)</b>				
(Potentially) pathogenic variant	Self-sample  (n=27)	Pap smear  (n=26)	Pipelle  (n=23)	Surgical sample  (n=29)
<i>TP53</i>	2 (7)	6 (19)	2 (9)	17 (59)
<i>PIK3CA</i>	3 (11)	5 (15)	4 (17)	6 (21)
<i>CTNNB1</i>	0	1 (4)	0	4 (14)
<i>PTEN</i>	0	1 (4)	1 (4)	3 (10)
<i>KRAS</i>	0	1 (4)	2 (9)	2 (7)
<i>ARID1A</i>	2 (7)	1 (4)	1 (4)	1 (3)
<i>MTOR</i>	0	0	0	1 (3)
<i>POLE</i>	0	0	0	0
Total number of pathogenic variants	7	15	10	34
Patients without a pathogenic variant	21 (78)	16 (62)	16 (70)	5 (17)
<b>Control patients (n=32)</b>				
(Potentially) pathogenic variant	Self-sample  (n=31)	Pap smear  (n=32)	Pipelle  (n=32)	Surgical sample  (n=31)
<i>TP53</i>	0	0	0	0
<i>PIK3CA</i>	0	2 (6)	2 (6)	0
<i>CTNNB1</i>	0	0	0	0
<i>PTEN</i>	0	0	0	0
<i>KRAS</i>	0	0	0	2 (6)
<i>ARID1A</i>	1 (3)	0	0	0
<i>MTOR</i>	0	0	0	0
<i>POLE</i>	0	0	0	0
Total number of pathogenic variants				
Patients without a pathogenic variant	30 (97)	30 (94)	30 (94)	29 (94)

Values are presented as number of pathogenic variants (% of patients). Percentages may total >100% as one patient can have multiple variants.

table 4). Two control patients (6%) had a *PIK3CA* variant which were not found in their other specimens.

### Diagnostic Accuracy

Among the 29 patients with ovarian cancer, at least one (potentially) pathogenic variant was detected in 83% of the surgical specimens, in 22% of the self-samples, in 38% of the Pap smears, and in 30% of the pipelles. In the patients with ovarian cancer the detection rate for a (potentially) pathogenic variant in any of the sampling methods (the self-sample, Pap smear, or pipelle) was 52%. Among the controls, a false positive variant was detected in 6%, 3%, 6%, and 6%, respectively, of the specimens. Detection rates were

roughly similar between early and late stage (see Supplementary Document 2). No correlation was found between the variant allele frequency in the ovarian tumor and the likelihood of detecting the variant in any of the sampling methods (data not shown).

The diagnostic accuracy of overlapping (potentially) pathogenic variants—for example, a pathogenic variant in minimally one of the sampling methods among patients with a pathogenic variant in their surgical specimen (n=24) and controls without a pathogenic variant in the surgical specimen (n=30)—is shown in Table 3. Sensitivity and specificity for an overlapping pathogenic variant in the surgical specimen and any of the sampling

**Table 3** Diagnostic accuracy of various combinations of measurements to detect ovarian cancer

	Sensitivity	Specificity
Any (potentially) pathogenic variant (all patients)		
	22 (9 to 43)	97 (83 to 100)
	38 (20 to 59)	94 (79 to 99)
	30 (13 to 53)	94 (79 to 99)
	52 (33 to 71)	88 (71 to 96)
(Potentially) pathogenic variants overlapping with the surgical specimen		
	5 (0 to 23)	97 (83 to 100)
	26 (10 to 48)	93 (80 to 99)
	6 (0 to 27)	93 (80 to 99)
	29 (13 to 51)	87 (69 to 96)
Diagnostic accuracy includes sensitivity and specificity. Values are presented as percentage (95% CI)		

methods were 29% (95% CI 13% to 51%) and 87% (95% CI 69% to 96%), respectively.

When analyzing *TP53* pathogenic variants only, we identified 21 patients with ovarian cancer with 22 *TP53* variants, of whom 17 had the *TP53* variant in their surgical specimen. Overlapping *TP53* variants with the surgical specimen were found in the Pap smear in three patients (sensitivity 18% (95% CI 4% to 43%)). No overlapping *TP53* variants were found in the self-samples or the pipelles. Among controls, no *TP53* variants were detected (100% specificity).

## DISCUSSION

### Summary of Main Results

In this multicenter prospective study we investigated the diagnostic accuracy of detecting ovarian cancer with mutational analysis in cervicovaginal and endometrial biopsies. We found (potentially) pathogenic variants in 83% of the ovarian cancer tumors, of which 71% were *TP53* variants. In 52%, a (potentially) pathogenic variant could be detected in either a cervicovaginal self-sample, Pap smear, or pipelle. Sensitivity was low for all sampling methods and remained low when analyzing *TP53* variants only. Among the controls, hardly any variants were detected, resulting in very high specificity of all sampling methods.

### Results in the Context of Published Literature

Despite an impressive research effort to improve the therapeutic options for ovarian cancer, the survival rate has barely increased over the past decades.<sup>2</sup> Diagnosing epithelial ovarian cancer in an early stage might improve prognosis substantially, emphasizing the need for early detection methods.<sup>2</sup> It is notable that early stage detection seemed not to be inferior to late stage detection in our study. Based on anatomical position, one could reason that pathogenic variants would be more frequently found in endometrial biopsies compared with cervicovaginal samples. Some studies reported on potential precursors of serous epithelial ovarian cancer in the endometrium,<sup>15,16</sup> although the fallopian tubes are nowadays considered as the site of origin of mainly serous epithelial ovarian cancer.<sup>4</sup> Thus far, no research has been published about sampling the endometrium with a pipelle biopsy to potentially detect ovarian cancer early. However, as we found, mutational analysis of the endometrium obtained via a pipelle seems not to be appropriate for this purpose, although this does not exclude the potential role of the uterus and/or endometrium in early ovarian cancer detection. The lower prevalence of diagnosed pathogenic variants in the pipelle biopsies compared with the cervicovaginal samples could be explained by the fact that endometrial tissue was processed in paraffin, in which DNA preservation is less optimal leading to a lower sequence coverage and thus lower sensitivity. The cervicovaginal samples were stored in PreservCyt medium, which better maintains DNA stability.<sup>17</sup> Future studies could investigate whether the detection rate of pathogenic variants would increase when analyzing DNA from pipelle samples being preserved in PreservCyt medium.

As demonstrated, cytology samples might be more promising in early ovarian cancer detection than endometrial histology samples. This is especially attractive as obtaining cervical cytology samples is highly accepted and less invasive than obtaining endometrial histology. Of the 34 detected variants in the surgical specimens, eight were also found in the Pap smear (24%) whereas we only detected one variant in the cytological cervicovaginal self-sample overlapping with the surgical specimen (3%). Our results are very similar to those of Wang et al,<sup>7</sup> who found that 29% of patients with ovarian cancer harbored detectable variants, mostly *TP53*, in their Pap smears. They also investigated intra-uterine cytology sampling using Tao brushes which could detect a variant in 42% of patients with ovarian cancer. Current research is investigating whether uterine cytology samples, obtained via lavage of the uterine cavity and analyzed with next-generation sequencing, can serve as an early detection method (NCT 02039388). Combining the results of their uterine cytology samples with cytologic assessed pipelle biopsies might show a new insight into the etiology of ovarian cancer. Also, the methylation status of such cytologic samples may contribute/improve ovarian cancer detection. Moreover, Barrett et al recently demonstrated that the DNA methylome in cervical samples can predict the risk of ovarian cancer with about 75% certainty.<sup>18</sup> Evaluation of the vaginal microbiome as a possible early detection method could also be promising. The microbiome might impact estrogen metabolism and may influence the risk of ovarian cancer, like exogenous estrogens.<sup>19–22</sup> In colorectal oncogenesis the microbiota seem to play a major role,<sup>23</sup> which may also apply to gynecological cancers.

## Original research

Considering the oncogenesis of high-grade serous carcinoma, we expected to find more *TP53* pathogenic variants. Moreover, approximately 90% of all patients with serous epithelial ovarian cancer have a *TP53* variant.<sup>14</sup> We found a *TP53* variant in 59% of patients, which might be explained by tumor heterogeneity.<sup>24</sup> Also, three of our five patients without a pathogenic variant in the surgical specimen had a high-grade serous carcinoma. Further, *ARID1A* variants were under-represented among patients with clear cell histology based on TCGA, probably because only four patients had (mixed) clear cell histology. Our detection rate may have been higher if we had included genes involved in homologous recombination as these are commonly related with epithelial ovarian cancer. For example, somatic *BReast CAncer* pathogenic variants can be detected in about 17% of all patients with epithelial ovarian cancer.<sup>25</sup>

### Strengths and Weaknesses

Our study is the first to sample the endometrium with a pipelle to potentially detect ovarian cancer. We covered most of the Müllerian tract by sampling the endometrium, cervix, and vagina, in addition to the tumor. There are, however, some limitations. A larger sample size would strengthen our results. The low prevalence of (potentially) pathogenic variants overall might be related to insufficiently deep sequencing of some samples, the choice of the library preparation method, and the content of the gene panel, although we expected that the majority would be picked up with this panel. The detection rate may have been higher when the pipelle samples were stored in PreservCyt medium. There might be some false positive variants as healthy persons appear sometimes to have pathogenic variants as well.<sup>26 27</sup> Furthermore, the class III pathogenic variants were considered (potentially) pathogenic, although these reflect a minority of all variants. Thus far it is unknown whether or not these variants should be considered pathogenic.

### Implications for Practice and Future Research

This study contributes to the unraveling of ovarian cancer etiology and assists in research regarding the urgently needed detection of ovarian cancer. For future research it would be relevant to investigate cytology samples acquired along the Müllerian tract, stored in PreservCyt medium, using alternative library preparation methods, expanding the gene panel, and potentially analyzing the methylation status or the vaginal microbiome.

### CONCLUSIONS

We investigated whether mutational analysis of samples along the Müllerian tract could be used to detect ovarian cancer. Diagnostic accuracy with our analysis was low for cervicovaginal self-samples, Pap smears, and endometrial biopsies when comparing the pathogenic variants in the samples to the variants in the tumor itself. Thus, these samples should not be used for (early) ovarian cancer detection.

#### Author affiliations

<sup>1</sup>Obstetrics & Gynaecology, Radboud Institute for Health Science, Radboudumc, Nijmegen, The Netherlands

<sup>2</sup>Pathology, Radboudumc, Nijmegen, The Netherlands

<sup>3</sup>Obstetrics and Gynaecology, Catharina Hospital, Eindhoven, The Netherlands

<sup>4</sup>Pathology, Catharina Hospital, Eindhoven, The Netherlands

<sup>5</sup>Obstetrics and Gynaecology, Elisabeth-TweeSteden Ziekenhuis, Tilburg, The Netherlands

<sup>6</sup>Human Genetics, Radboudumc, Nijmegen, The Netherlands

<sup>7</sup>Radiation Oncology, Radboudumc, Nijmegen, The Netherlands

**Contributors** MvB wrote the main manuscript text. MvB, LvdP, MS, HKvdV, ML, AE, JdH, JP, and CR provided the conceptualization and methodology of the study. MvB, LvdP, JB, MS, SS, ML, AE, JdH, JP, and CR were involved in the analysis of the data. CR acted as guarantor. All authors reviewed the manuscript.

**Funding** The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

**Competing interests** None declared.

**Patient consent for publication** Consent obtained directly from patient(s)

**Ethics approval** This study involves human participants and was approved by the Medical-Ethical Committee of Arnhem-Nijmegen study number 2013/451. Participants gave informed consent to participate in the study before taking part.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Data are available upon reasonable request. All data relevant to the study are included in the article or uploaded as supplementary information. The data supporting the conclusions of this article are available upon reasonable request to the corresponding author. In accordance with the journal's guidelines, we will provide our data for independent analysis by a selected team by the Editorial Team for the purposes of additional data analysis or for the reproducibility of this study in other centers if such is requested.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution 4.0 Unported (CC BY 4.0) license, which permits others to copy, redistribute, remix, transform and build upon this work for any purpose, provided the original work is properly cited, a link to the licence is given, and indication of whether changes were made. See: <https://creativecommons.org/licenses/by/4.0/>.

#### ORCID iDs

Majke H.D. van Bommel <http://orcid.org/0000-0002-7562-9861>

Heidi V.N. Küsters-Vandevelde <http://orcid.org/0000-0002-1012-941X>

#### REFERENCES

- Sung H, Ferlay J, Siegel RL, *et al.* Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2021;71:209–49.
- Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. *CA Cancer J Clin* 2019;69:7–34.
- Menon U, Gentry-Maharaj A, Burnell M, *et al.* Ovarian cancer population screening and mortality after long-term follow-up in the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS): a randomised controlled trial. *Lancet* 2021;397:2182–93.
- Labidi-Galy SI, Papp E, Hallberg D, *et al.* High grade serous ovarian carcinomas originate in the fallopian tube. *Nat Commun* 2017;8:1093.
- Hermesen BBJ, Olivier RI, Verheijen RHM, *et al.* No efficacy of annual gynaecological screening in BRCA1/2 mutation carriers; an observational follow-up study. *Br J Cancer* 2007;96:1335–42.
- Kinde I, Bettgowda C, Wang Y, *et al.* Evaluation of DNA from the Papanicolaou test to detect ovarian and endometrial cancers. *Sci Transl Med* 2013;5:ra4.
- Wang Y, Li L, Douville C, *et al.* Evaluation of liquid from the Papanicolaou test and other liquid biopsies for the detection of endometrial and ovarian cancers. *Sci Transl Med* 2018;10:eap8793. doi:10.1126/scitranslmed.aap8793
- Maritschnegg E, Wang Y, Pecha N, *et al.* Lavage of the uterine cavity for molecular detection of Müllerian duct carcinomas: a proof-of-concept study. *J Clin Oncol* 2015;33:4293–300.

- 9 Reijnen C, van der Putten LJM, Bulten J, *et al.* Mutational analysis of cervical cytology improves diagnosis of endometrial cancer: a prospective multicentre cohort study. *Int J Cancer* 2020;146:2628–35.
- 10 Eijkelenboom A, Kamping EJ, Kastner-van Raaij AW, *et al.* Reliable next-generation sequencing of formalin-fixed, paraffin-embedded tissue using single molecule tags. *J Mol Diagn* 2016;18:851–63.
- 11 Cancer Genome Atlas Research Network, Kandoth C, Schultz N, *et al.* Integrated genomic characterization of endometrial carcinoma. *Nature* 2013;497:67–73.
- 12 Cancer Genome Atlas Research Network. Integrated genomic analyses of ovarian carcinoma. *Nature* 2011;474:609–15.
- 13 cfCG. Ovarian Epithelial Tumor - Serous Ovarian Cancer, 2014. Available: <http://www.cbiportal.org/>
- 14 Ghezelayagh TS, Pennington KP, Norquist BM, *et al.* Characterizing TP53 mutations in ovarian carcinomas with and without concurrent BRCA1 or BRCA2 mutations. *Gynecol Oncol* 2021;160:786–92.
- 15 Massuger L, Roelofsen T, van Ham M. The origin of serous ovarian cancer may be found in the uterus: a novel hypothesis. *Med Hypotheses* 2010;74:859–61.
- 16 Roelofsen T, van Kempen LCLT, van der Laak JAWM, *et al.* Concurrent endometrial intraepithelial carcinoma (EIC) and serous ovarian cancer: can EIC be seen as the precursor lesion? *Int J Gynecol Cancer* 2012;22:457–64.
- 17 Filho AL, Gonçalves AEP, Martinho O, *et al.* Liquid-based cytology in DNA-based molecular research: viability and potential application. *Anal Quant Cytol Histol* 2009;31:395–400.
- 18 Barrett JE, Jones A, Evans I, *et al.* The DNA methylome of cervical cells can predict the presence of ovarian cancer. *Nat Commun* 2022;13:448.
- 19 Havrilesky LJ, Moorman PG, Lowery WJ, *et al.* Oral contraceptive pills as primary prevention for ovarian cancer: a systematic review and meta-analysis. *Obstet Gynecol* 2013;122:139–47.
- 20 Buchta V. Vaginal microbiome. *Ceska Gynekol* 2018;83:371–9.
- 21 Chambers LM, Bussies P, Vargas R, *et al.* The microbiome and gynecologic cancer: current evidence and future opportunities. *Curr Oncol Rep* 2021;23:92.
- 22 Néné NR, Reisel D, Leimbach A, *et al.* Association between the cervicovaginal microbiome, BRCA1 mutation status, and risk of ovarian cancer: a case-control study. *Lancet Oncol* 2019;20:1171–82.
- 23 Abu-Ghazaleh N, Chua WJ, Gopalan V. Intestinal microbiota and its association with colon cancer and red/processed meat consumption. *J Gastroenterol Hepatol* 2021;36:75–88.
- 24 Erickson BK, Kinde I, Dobbin ZC, *et al.* Detection of somatic TP53 mutations in tampons of patients with high-grade serous ovarian cancer. *Obstet Gynecol* 2014;124:881–5.
- 25 Vos JR, Fakkert IE, de Hullu JA, *et al.* Universal tumor DNA BRCA1/2 testing of ovarian cancer: prescreening PARPi treatment and genetic predisposition. *J Natl Cancer Inst* 2020;112:161–9.
- 26 Maritschnegg E, Heitz F, Pecha N, *et al.* Uterine and tubal lavage for earlier cancer detection using an innovative catheter: a feasibility and safety study. *Int J Gynecol Cancer* 2018;28:1692–8.
- 27 Moore L, Leongamornlert D, Coorens THH, *et al.* The mutational landscape of normal human endometrial epithelium. *Nature* 2020;580:640–6.



## Supplementary

Supplementary Document 1. Workflow for analyzing pathogenic variants	p. 2
Supplementary Table 1. Gene regions targeted in our single molecule molecular inversion probes panel	p. 5
Supplementary Figure 1. Mean numbers of unique reads for each variant region	p. 6
Supplementary Table 2. Coverage details per sequenced exon	p. 8
Supplementary Table 3. Baseline characteristics of each patient	p. 11
Supplementary Table 4. Identified pathogenic variants in the four specimens per patient	p. 14
Supplementary Document 2. Detection rate per sampling method by stage	p. 17

**Supplementary Document 1. Workflow for analyzing pathogenic variants****Tissue processing**

The cervicovaginal self-sample (Evalyn Brush, Rovers Medical Devices B.V., Oss, the Netherlands) was taken according to the manufacturer's instructions. The brush tips of the cervicovaginal self-sample and Pap smear were suspended in Preservcyt medium (Hologic, Marlborough, MA). The brush tips of the cervicovaginal self-sample were suspended in 4.5ml vials, which were vortexed for 3 x 15s, stored overnight at 4 °C, and again vortexed for 2 x 15s, before the brushes were removed and discarded. The samples were stored at room temperature until DNA extraction. A representative Formalin-Fixed Paraffin-Embedded surgical sample of the ovary was selected at the end of the surgery in both ovarian cancer patients and in women undergoing surgery for a benign gynecological condition (control group).

**DNA Extraction**

Representative areas of the ovarian cancer in the surgical specimens were marked and selected by macrodissection from 2 x 20 µm thick FFPE sections (C.R., J.B.). From the pipelle endometrial biopsies, two 20 µm thick sections were fully used for DNA extraction, without selecting tissue by macrodissection. The cervicovaginal self-samples and Pap smears were centrifuged for 5 minutes at 14.000 x g, after which the supernatant was transferred and the pellets were centrifuged for 1 minute at 14.000 x g. The remaining pellet was used for DNA extraction. These specimens were digested at 56°C overnight in TET-lysisbuffer (10mmol/L Tris/HCL pH8.5, 1mmol/L EDTA pH8.0, 0.01% Tween-20) with 5% Chelex-100 (Bio-Rad, Hercules, CA, US) and 0.2% proteinase K, followed by inactivation at 95°C for ten minutes. Subsequently, the supernatant was transferred after centrifugation into a clean tube. DNA concentration was determined using the Qubit Broad Range Kit (Thermo Fisher Scientific, Waltham, MA, US).

**Single molecule molecular inversion probes panel design and library preparation**

The samples were analyzed using single molecule Molecular Inversion Probes (Integrated DNA Technologies, Leuven, Belgium). The design of the single molecule Molecular Inversion Probes and the library preparation were performed as previously described [10]. Briefly, single molecule Molecular Inversion Probes were designed in a tiling manner for all included hotspots in oncogenes relevant in endometrial and ovarian cancer (*CTNNB1*, *KRAS*, *MTOR*, *PIK3CA* and *POLE*) and all coding and splice site sequences of tumor suppressor genes (*ARID1A*, *PTEN*, and *TP53*, **Supplementary Table**

1) preferentially targeting both strands with two independent single molecule Molecular Inversion Probes (probe sequences available on request). The single molecule Molecular Inversion Probes consisted of an extension and ligation probe arm, together 40bp long, with a gap of 112 bp, with a common backbone sequence for PCR-based library amplification. The ligation probe arm and backbone are connected with a backbone, also containing an 8bp degenerate sequence (8xN) serving as a Unique Molecular Identifier (UMI, also known as 'single molecule tag'). The single molecule Molecular Inversion Probes were mixed and phosphorylated using 1  $\mu$ L of T4 polynucleotide kinase (M0201; New England Biolabs, Ipswich, MA, US) per 25  $\mu$ L of 100  $\mu$ mol/L smMIPs and ATP-containing G4 DNA ligase buffer (B0202, New England Biolabs). The molecular ratio between gDNA and single molecule Molecular Inversion Probes was 1:3200 for every individual smMIP and the standard genomic DNA input was set at 100 ng.

A capture mix was made (total capture volume 25  $\mu$ L) containing the phosphorylated smMIP pool, 1 unit of Ampligase DNA ligase (A0110K; EpiBio, Madison, WI) and Ampligase Buffer (A1905B, DNA ligase buffer), 3.2 units of Hemo Klentaq (M0332; New England Biolabs), finally 8 mmol of dNTPs (28-4065-20/-12/-22/-32; GE Healthcare, Little Chalfont, UK) and, when available, 100 ng of genomic DNA in a 20  $\mu$ L volume. Subsequently, this capture mix was denatured (95°C for 10 minutes) and incubated for probe hybridization, extension and ligation at 60°C for 18 hours. After cooling, exonuclease treatment was performed by adding Exonuclease I (10 units; M0293; New England Biolabs) and III (50 units; M0206; New England Biolabs) and Ampligase Buffer to the capture mix (total of 27  $\mu$ L) and incubating at 37°C for 45 minutes, with subsequent inactivation at 95°C for 2 minutes. A total of 20  $\mu$ L was used for PCR in a total volume of 50  $\mu$ L including a common forward primer, bar-coded reverse primers, and iProof high fidelity master mix (1725310, Bio-Rad, Venendaal, the Netherlands). The resulting PCR products were pooled prior to purification with 0.8x volume of Agencourt Ampure XP Beads (A63881, Beckman Coulter, Woerden, the Netherlands). Each cytology sample was assessed in two independent library preparations, because of the expected low mutant allele frequencies.

### Sequencing and analysis

Sequencing of the purified libraries, denatured and diluted to 1.2pmol/L, was performed on a NexSeq500 device (Illumina, San Diego, CA, US) according to the manufacturer's instructions (300 cycles High Output sequencing kit, v2), resulting in 2x150bp paired-end reads. The resulting Bcl files were converted to fastq files and bar-coded reads were subsequently demultiplexed. Single-molecule-directed assembly of duplicate reads was performed to generate consensus ('unique') reads using the analysis software Sequence Pilot (version 4.4.0; JSI medical systems, Ettenheim,

Germany). For variant calling in Sequence Pilot, variant detection thresholds were generally set to 3% (surgical specimens and pipelles) or 1% (Pap smears and self-samples) of all unique reads at that position and a minimum of 5 unique reads representing  $\geq 3$  individual gDNA molecules.

Variants were classified as “pathogenic”, “likely pathogenic”, “variant of unknown significance”, “likely benign” and “benign” and the first three categories were considered (potentially) pathogenic. Synonymous variants were only considered when present at exon ends. Finally, intronic variants were excluded with the exception of splice site sequences. First, the surgical specimens (of ovarian cancer or a benign gynecological condition) were analyzed for the presence of somatic variants using variant calling. Subsequently, these variants were investigated in pipelle endometrial biopsies and the cytology samples, using a three-step method. First, variants yielded by variant calling were assessed. Second, all samples were manually screened for the presence of reads harboring the known variants. In case no variant was found in the surgical specimen, the cytology samples were searched for the presence of known hotspot variants in *CTNNB1*, *KRAS*, *MTOR*, *PIK3CA* and *POLE*.

After this, assessment of variants in the pipelle endometrial biopsies and cytology samples was performed, with adjusted variant calling settings for hotspots in oncogenes (minimum of 5 unique reads representing 3 gDNA molecules, without minimal variant allele frequencies). Settings for tumor suppressor genes were unadjusted.

To determine whether sufficient DNA molecules were sequenced to reliably (>95% certainty) exclude variants above a certain mutant allele frequency, a cumulative binomial distribution was used that calculated the required unique read depths [10]. For all surgical specimens and pipelle endometrial biopsies these required read depths were assessed in the context of the estimated tumor load (percentage of neoplastic cells). For all cytology samples, a unique read depth representing > 250 individual gDNA molecules at each hotspot position was pursued, in order to reduce the chance to less than 5% of missing variants with an variant allele frequency above 3%. In case hotspot positions were sequenced with insufficient unique read depth, library preparation and sequencing was repeated (once or twice if needed). Mean numbers of unique reads for each variant region are shown in **Supplementary Figure 1**.

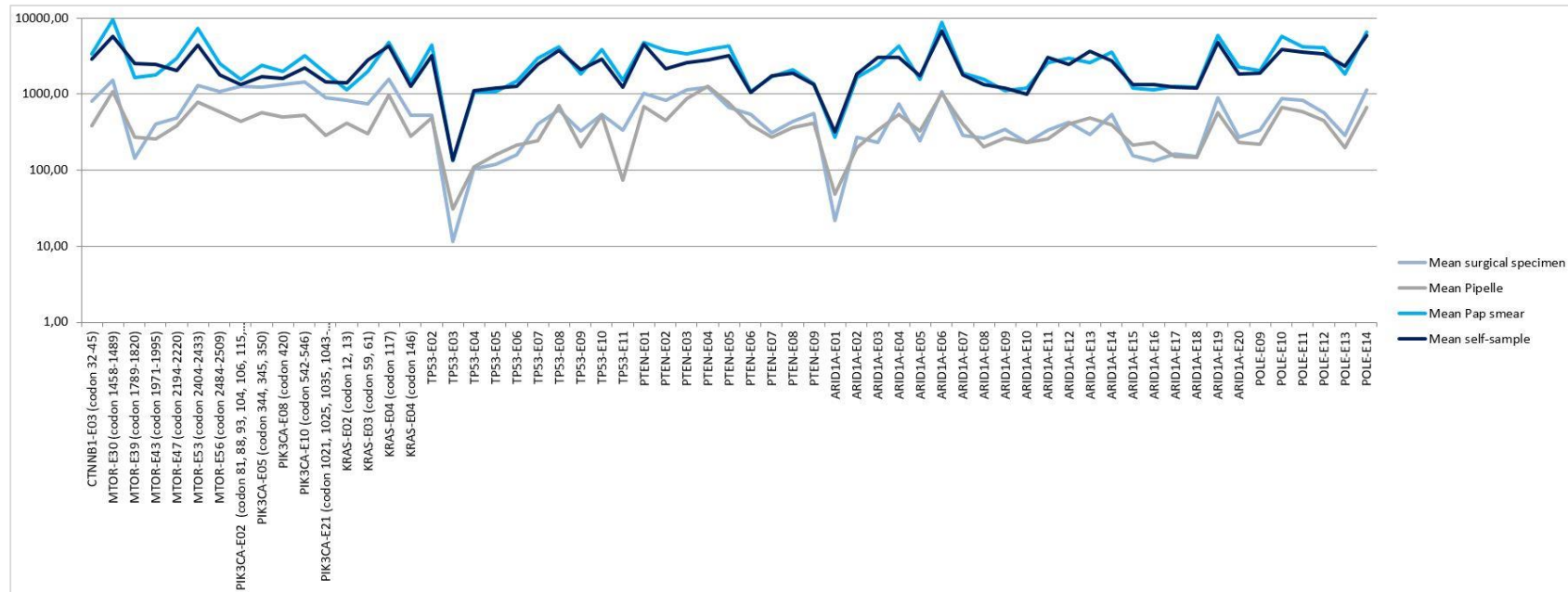
In order to exclude potential false positive calls due to PCR, sequencing, or Formalin-Fixed Paraffin-Embedded deamination artifacts, the background was determined at all sequenced positions using 12 control Pap smears and cervicovaginal self-samples. Because proliferating endometrium could harbor *PIK3CA* and *KRAS* mutations, only cytology samples from patients with histo-pathologically proven atrophic endometrium were selected. The variant allele frequency to background (signal to noise) ratio was > 5 for all identified variants.

**Supplementary Table 2. Gene regions targeted in our single molecule Molecular Inversion Probes panel**

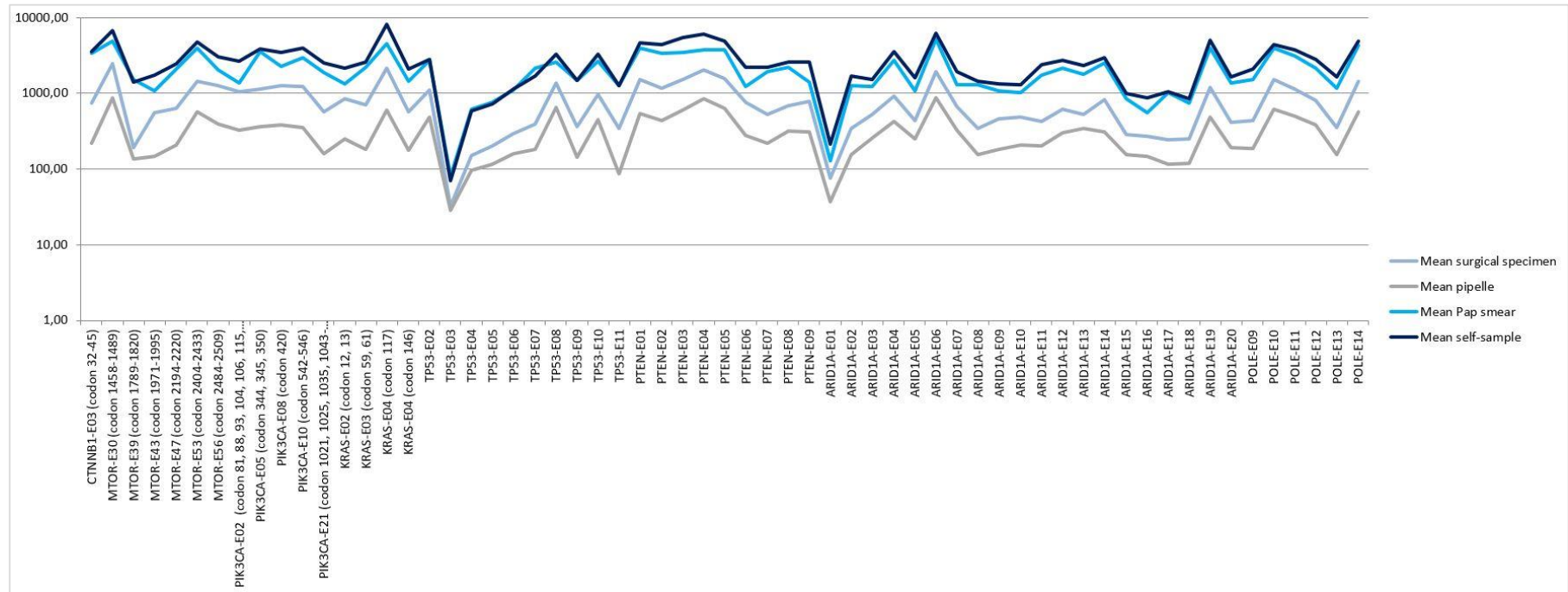
Gene	Exon	Targeted codons	Positions	RefSeq ID	Ensembl ID
<i>ARID1A</i>	1 to 20	M1-Stop2286	c.1 to c.6858	NM_006015.5	ENST00000324856
<i>CTNNB1</i>	3	D32-S45	c.53 to c.146	NM_001904.3	ENST00000349496
<i>KRAS</i>	2	G12-G13	c.9 to c.71	NM_004985.4	ENST00000311936
	3	A59-Q61	c.122 to c.215		
	4	K117, A146	c.291-5 to c.357		
			c.402 to c.450+5		
<i>MTOR</i>	30	D1458-E1489	c.4371 to c.4469+5	NM_004958.3	ENST00000361445
	39	A1789-A1820	c.5365-5 to c.5460		
	43	A1971-L1995	c.5911-5 to c.5985		
	47	Q2194-L2220	c.6580 to c.6662+5		
	53	M2404-D243	c.7210 to c.7300+5		
	56	G2484-T2509	c.7448-5 to c.7527		
<i>PIK3CA</i>	10	E542-Q546	c.1558 to c.1664+5	NM_006218.3	ENST00000263967
	21	M1043-G1049	c.3058 to c.3207+10		
<i>POLE</i>	9 to 14	D268-E491	c.802-5 to c.1473+5	NM_006231.3	ENST00000320574
<i>PTEN</i>	1 to 9	M1-Stop404	c.1 to c.1210+5	NM_000314.6	ENST00000371953
<i>TP53</i>	2 to 11	>95% of all coding and splice sequences (-5/+5)	c.1 to c.1180+5	NM_000565.5	ENST00000269305

Supplementary Figure 1. Mean numbers of unique reads for each variant region

## A. Ovarian cancer patients (n=29)



## B. Control patients (n=32)



Supplementary Table 2. Coverage details per sequenced exon

Exon	Ovarian cancer patients												Control patients											
	Self-sample			Pap smear			Pipelle			Surgical specimen			Self-sample			Pap smear			Pipelle			Surgical specimen		
	Mean	P25	P75	Mean	P25	P75	Mean	P25	P75	Mean	P25	P75	Mean	P25	P75	Mean	P25	P75	Mean	P25	P75	Mean	P25	P75
CTNNB1-E03 (codon 32-45)	2861.5	872.5	4443.5	3408.4	1854.0	4575.0	387.0	35.0	562.5	796.6	122.5	1139.3	3597.7	1755.3	4332.3	3381.1	1578.0	4083.0	222.2	36.0	346.0	753.4	408.0	786.0
MTOR-E30 (codon 1458-1489)	5759.2	1361.0	7908.0	9585.1	2979.0	15141.0	1068.9	96.0	1523.0	1536.1	420.0	1979.8	6746.8	2850.0	7926.5	4944.3	2660.0	6212.0	866.0	142.0	1322.0	2497.9	1486.0	2946.0
MTOR-E39 (codon 1789-1820)	2510.0	416.0	3382.0	1658.7	800.0	2288.0	274.9	13.0	412.0	145.4	24.5	210.0	1416.7	611.5	2032.3	1518.9	557.0	1908.0	137.7	16.0	186.0	191.8	72.0	180.0
MTOR-E43 (codon 1971-1995)	2476.8	1136.5	3708.0	1805.4	865.0	2678.0	255.1	17.0	364.0	400.2	83.8	573.3	1727.1	988.0	2028.5	1081.7	556.0	1610.0	146.5	20.0	214.0	551.0	352.0	682.0
MTOR-E47 (codon 2194-2220)	2037.1	722.5	3208.0	2955.1	1344.0	3986.0	380.1	24.0	537.0	491.4	107.5	609.0	2488.6	1344.5	2917.5	2116.6	1239.0	2591.0	208.1	56.0	308.0	640.9	368.0	689.0
MTOR-E53 (codon 2404-2433)	4373.2	998.0	5537.5	7281.9	1604.0	12030.0	786.3	64.0	1114.0	1294.8	316.0	1494.5	4831.7	1634.5	6646.0	3940.9	2118.0	5273.0	576.1	148.0	838.0	1428.6	972.0	1950.0
MTOR-E56 (codon 2484-2509)	1801.1	434.0	2466.0	2542.4	675.0	3796.0	585.1	47.0	754.0	1084.8	202.0	1339.0	3067.8	1312.5	4230.8	2043.2	620.0	3022.0	393.8	40.0	660.0	1272.3	762.0	1724.0
PIK3CA-E02 (codon 81, 88, 93, 104, 106, 115, 118)	1326.9	378.0	1886.0	1574.0	300.0	1930.0	435.3	24.0	620.0	1278.4	278.0	1355.0	2637.3	1007.0	3315.0	1353.8	480.0	1906.0	331.0	32.0	558.0	1059.9	684.0	1396.0
PIK3CA-E05 (codon 344, 345, 350)	1715.9	895.5	2369.0	2428.5	862.0	3204.0	571.1	44.0	776.0	1247.0	372.0	1217.8	3854.3	2400.0	5360.3	3527.1	1926.0	5059.0	368.0	54.0	580.0	1130.5	714.0	1284.0
PIK3CA-E08 (codon 420)	1611.9	400.5	2536.5	2002.8	570.0	2554.0	495.0	38.0	692.0	1342.9	362.0	1417.5	3455.3	1652.8	4912.3	2291.7	796.0	3704.0	383.7	28.0	652.0	1254.2	790.0	1674.0
PIK3CA-E10 (codon 542-546)	2219.6	705.0	3184.0	3245.4	1293.0	4495.0	523.5	47.0	739.0	1445.8	339.8	1409.0	3963.1	1901.0	5747.0	2974.8	1424.0	3986.0	352.2	44.0	590.0	1235.1	741.0	1583.0
PIK3CA-E21 (codon 1021, 1025, 1035, 1043-1049, 1069)	1465.3	564.0	2378.5	1906.0	1072.0	2696.0	290.1	30.0	420.0	904.0	190.5	965.0	2513.1	1532.3	3545.8	1897.0	1037.0	2356.0	161.1	18.0	260.0	567.4	340.0	680.0
KRAS-E02 (codon 12, 13)	1398.0	301.0	1808.0	1147.6	130.0	1722.0	418.7	29.0	586.0	822.1	151.0	1072.0	2147.2	833.0	3126.5	1322.5	312.0	2101.0	252.6	18.0	418.0	860.4	554.0	1182.0
KRAS-E03 (codon 59, 61)	2804.1	1264.5	4135.5	1996.5	1118.0	2744.0	304.0	38.0	449.0	746.0	107.5	1095.0	2599.2	1447.8	3718.3	2218.3	1004.0	3010.0	184.3	18.0	264.0	715.0	420.0	858.0
KRAS-E04 (codon 117)	4252.0	1758.5	6797.0	4795.2	2433.0	5970.0	971.7	77.0	1248.0	1557.4	438.0	2006.5	8186.1	4196.5	11352.5	4554.1	2827.0	6478.0	609.5	64.0	952.0	2163.5	1376.0	2549.0
KRAS-E04 (codon 146)	1274.2	367.0	1889.5	1441.8	398.0	2242.0	282.2	19.0	399.0	523.6	133.5	645.5	2123.8	946.8	3088.5	1456.4	446.0	2176.0	176.1	16.0	290.0	577.9	370.0	757.0
TP53-E02	3190.1	1246.7	5058.8	4406.8	3314.6	5814.9	490.6	59.8	850.7	531.1	128.6	532.4	2802.1	1283.6	4023.1	2705.8	1475.4	3187.6	491.5	114.0	591.0	1105.1	550.0	1370.5
TP53-E03	136.9	60.0	155.0	133.7	56.0	202.0	31.2	2.0	46.0	11.6	2.0	14.0	70.5	35.5	101.5	79.6	58.0	94.0	28.6	6.0	46.0	32.6	14.0	46.0
TP53-E04	1101.2	514.9	1679.7	1056.9	699.1	1469.5	111.4	15.0	194.9	105.1	29.0	128.8	584.0	291.1	827.7	623.0	346.9	830.4	96.5	21.7	126.2	153.1	99.0	205.0
TP53-E05	1212.3	363.7	1658.7	1080.1	353.3	1626.2	159.6	22.4	257.5	118.0	25.8	157.5	732.8	347.2	1014.2	764.0	348.1	1064.6	114.8	37.3	154.8	201.3	96.6	205.1



TP53-E06	1282.0	543.9	1960.0	1498.4	879.0	2173.8	214.8	28.8	373.6	157.9	33.9	173.0	1143.6	586.2	1475.0	1123.5	597.4	1456.4	159.5	35.3	223.6	294.7	161.5	284.6
TP53-E07	2464.8	1134.9	3813.8	2977.0	1745.9	3786.5	243.9	27.5	387.2	405.2	94.0	369.0	1673.7	741.8	2148.1	2163.2	756.2	2815.6	180.7	43.2	247.2	391.8	187.3	503.8
TP53-E08	3752.0	1730.3	5399.1	4196.0	2099.1	5729.4	708.0	73.5	1208.4	615.0	145.6	766.7	3267.1	1934.8	4527.0	2577.8	1624.3	3195.9	648.3	174.7	862.9	1377.4	844.4	1617.8
TP53-E09	2081.6	883.4	2867.1	1837.4	1077.7	2666.9	200.8	18.8	336.2	326.3	75.6	397.7	1489.6	834.8	2025.0	1472.2	765.4	2088.4	143.0	31.0	181.6	366.5	230.9	457.7
TP53-E10	2889.5	784.6	4786.0	3884.5	2196.3	4888.3	529.6	46.5	875.9	542.0	141.7	707.6	3295.2	1697.9	4373.0	2645.0	1515.9	3412.9	454.9	111.1	547.8	971.5	539.6	1050.4
TP53-E11	1233.2	622.6	1493.9	1508.4	972.7	2215.9	73.8	12.3	106.9	334.9	103.0	462.8	1264.2	757.8	1641.1	1266.9	633.3	1451.0	87.1	15.4	83.2	343.1	155.9	388.9
PTEN-E01	4572.0	2469.6	6255.9	4722.7	3209.1	6046.1	693.4	86.7	1149.8	1019.7	249.2	1286.0	4715.7	3053.0	5921.6	3966.5	2577.1	5144.4	545.6	111.3	652.4	1530.2	964.9	1681.8
PTEN-E02	2129.7	1161.6	2947.7	3737.4	2058.7	4881.2	453.3	37.8	645.8	835.0	205.2	1047.9	4465.2	2460.5	6160.4	3353.6	1609.8	4641.8	438.6	60.9	569.7	1155.6	661.7	1289.4
PTEN-E03	2583.2	788.0	3736.5	3419.2	747.0	5184.0	871.7	46.0	1196.0	1130.8	334.0	1641.8	5496.7	2404.5	7613.8	3480.0	935.0	5685.0	601.9	64.0	826.0	1505.5	916.0	1882.0
PTEN-E04	2811.1	905.5	3619.0	3859.6	894.0	5950.0	1254.7	80.0	1943.0	1250.0	399.5	1732.5	6036.4	2813.0	9348.8	3774.7	1022.0	6208.0	843.5	126.0	1164.0	2053.9	1020.0	2343.0
PTEN-E05	3174.1	1502.1	5304.4	4253.1	2354.9	5383.2	768.4	61.8	1167.2	667.5	234.7	755.6	4844.3	2597.3	6325.7	3766.8	2176.6	4769.4	635.1	124.9	836.7	1584.5	916.0	1602.4
PTEN-E06	1049.5	428.0	1390.1	1075.7	350.4	1429.2	389.2	25.8	573.5	537.1	137.0	666.9	2223.6	1302.4	2884.1	1230.7	460.8	1747.7	278.4	21.8	411.1	765.8	313.9	905.9
PTEN-E07	1741.5	836.3	2614.2	1707.5	1173.7	2054.1	273.2	32.9	390.3	310.9	105.6	404.1	2208.3	1215.4	2945.4	1937.8	1129.0	2130.9	217.6	27.2	253.9	520.1	308.5	590.4
PTEN-E08	1864.1	946.7	2601.4	2121.2	1367.5	2787.4	367.8	38.1	534.0	438.2	162.7	558.3	2589.3	1542.5	3434.7	2188.7	1326.5	2780.6	313.9	50.0	440.2	691.8	425.0	853.3
PTEN-E09	1341.7	433.5	1822.7	1372.7	428.6	2171.5	413.0	30.1	546.6	557.9	180.8	736.1	2610.7	1329.3	3459.5	1423.6	773.7	1963.3	305.8	41.1	399.3	779.3	477.1	888.8
ARID1A-E01	318.8	113.9	420.9	271.6	126.9	400.6	48.9	5.8	61.8	21.9	7.2	36.5	214.2	109.5	253.8	129.8	62.6	162.0	37.3	12.9	50.8	75.8	38.8	86.1
ARID1A-E02	1838.2	600.1	2835.2	1671.7	1203.1	2146.9	199.7	18.5	301.2	271.7	73.9	362.6	1694.0	1034.1	2050.4	1273.5	703.7	1595.6	154.8	37.2	175.0	342.0	220.2	390.8
ARID1A-E03	3052.3	1451.8	4105.8	2424.9	1312.6	3569.5	334.8	44.2	569.7	234.5	58.5	261.1	1524.3	798.5	2124.7	1240.4	742.5	1725.2	254.0	71.7	336.0	524.0	297.5	537.6
ARID1A-E04	3021.1	1131.3	4525.8	4336.7	2172.9	6045.6	547.0	55.6	812.1	751.7	168.7	952.5	3574.4	1968.8	4664.6	2772.1	1666.7	3833.6	430.7	72.3	479.0	917.6	585.2	1114.9
ARID1A-E05	1759.3	669.4	2763.3	1570.6	994.9	2118.5	328.5	30.1	529.9	245.7	57.8	304.2	1597.9	959.3	2160.3	1093.7	609.8	1518.9	250.3	58.7	305.7	437.6	255.7	481.5
ARID1A-E06	6811.6	2524.2	10819.3	8718.4	4741.8	13472.0	1036.0	107.2	1634.0	1070.2	269.3	1372.0	6204.8	3765.6	7308.9	5193.8	2945.4	6536.8	873.6	238.0	1105.8	1927.9	1082.9	2295.2
ARID1A-E07	1785.2	702.0	2439.4	1884.9	950.9	2439.8	408.9	39.3	671.6	286.1	85.7	398.2	1917.1	1165.7	2673.5	1296.0	722.2	1670.5	323.9	81.7	440.5	672.6	448.1	718.2
ARID1A-E08	1346.9	483.3	2144.2	1582.8	1187.7	2043.8	203.9	20.7	315.3	264.3	68.8	334.7	1449.5	813.5	1829.3	1302.6	676.1	1694.2	155.5	30.1	194.0	346.0	201.1	393.9
ARID1A-E09	1209.3	500.2	1556.2	1121.0	708.2	1540.9	262.6	24.9	406.3	342.4	81.9	478.6	1340.9	713.1	1953.5	1085.4	655.3	1402.6	183.0	25.4	194.9	455.4	280.7	550.2
ARID1A-E10	1007.5	598.5	1256.7	1214.9	774.2	1549.8	230.2	25.3	386.1	230.7	59.5	285.2	1301.1	808.1	1799.1	1022.5	701.9	1375.6	207.8	44.0	296.2	491.6	283.8	589.3
ARID1A-E11	3023.7	1074.7	4673.5	2614.6	1583.1	3535.4	257.0	28.0	366.6	333.3	78.0	442.4	2426.7	1379.7	3051.8	1737.4	940.8	2174.3	200.9	45.2	230.9	430.4	242.6	474.8
ARID1A-E12	2464.6	1164.9	4001.5	2979.6	1939.6	4023.9	406.8	44.1	642.1	429.8	135.5	535.1	2720.7	1472.8	3539.4	2132.9	1322.8	2755.1	305.6	57.4	356.1	623.0	370.8	687.7
ARID1A-E13	3641.2	1567.0	5298.0	2574.0	1106.5	3911.0	487.2	58.6	820.3	291.8	73.5	417.6	2330.4	1134.6	3054.7	1775.4	1032.7	2309.9	341.3	80.8	440.2	533.5	249.8	636.1
ARID1A-E14	2765.4	1231.4	4329.0	3550.6	2651.0	4769.6	396.7	52.3	599.0	541.8	119.3	672.4	2951.6	1662.2	3895.9	2554.1	1544.5	3357.3	313.7	73.1	390.3	824.1	476.5	785.4
ARID1A-E15	1352.7	453.0	2121.9	1197.1	632.9	1617.2	211.6	21.6	355.3	154.0	41.6	216.2	1001.7	449.4	1294.1	838.9	410.2	1162.9	153.3	28.6	214.7	288.2	166.0	273.7

ARID1A-E16	1324.4	608.6	1957.8	1155.0	728.7	1899.8	232.8	22.7	402.3	134.0	31.8	195.1	871.6	473.4	1162.0	561.3	328.2	770.3	147.9	41.4	220.2	274.1	161.5	289.9
ARID1A-E17	1222.8	426.0	2017.8	1259.3	635.8	1739.5	153.3	16.6	240.5	165.5	34.4	221.3	1039.8	480.2	1568.9	1011.6	450.0	1272.7	116.7	27.9	160.1	243.5	131.4	247.6
ARID1A-E18	1203.3	440.1	1812.9	1221.1	722.2	1549.9	147.7	19.1	250.7	153.3	39.7	184.4	841.7	505.3	1075.0	739.7	512.4	923.5	120.1	28.4	157.1	247.6	147.1	272.8
ARID1A-E19	4809.4	1504.9	7174.3	5997.1	2456.5	8553.2	574.8	57.4	860.7	901.3	181.2	1035.9	4990.0	2897.5	6147.3	3921.6	2066.7	5136.4	483.6	82.5	541.2	1215.3	654.2	1141.3
ARID1A-E20	1859.0	675.1	3026.4	2257.9	1677.6	3086.6	229.9	30.2	360.1	272.3	70.3	320.8	1669.3	909.5	1986.1	1387.5	829.4	1789.9	193.3	42.8	232.8	418.1	258.8	419.8
POLE-E09	1873.4	818.8	2893.4	2036.2	1559.1	2915.9	221.8	18.4	355.6	338.8	97.8	415.9	2098.9	1296.7	2729.6	1525.2	845.6	1913.0	184.6	30.6	208.8	435.7	239.0	514.7
POLE-E10	3853.3	1564.1	5809.0	5764.9	2816.0	8392.7	673.8	62.7	1104.9	874.2	230.6	1374.3	4446.4	2563.0	5797.0	4015.1	2128.0	4787.3	612.7	133.7	768.9	1508.8	864.8	1726.7
POLE-E11	3543.9	1289.3	5650.8	4197.8	2087.1	6443.6	585.9	71.4	895.7	828.0	191.2	1188.2	3808.2	2327.9	4546.5	3127.6	1740.5	4133.1	504.6	103.4	544.9	1125.0	634.1	1431.1
POLE-E12	3399.6	883.6	5971.3	4066.8	2167.3	6869.9	449.9	54.6	725.1	571.2	131.2	834.8	2841.2	1251.4	3479.2	2147.4	1250.0	2706.8	386.2	86.9	451.8	813.1	475.9	975.3
POLE-E13	2362.9	919.8	2892.7	1845.1	1008.5	2892.4	196.7	17.0	316.4	285.5	67.5	361.5	1659.6	950.3	1728.5	1168.3	631.7	1427.4	155.4	42.8	158.6	349.2	220.3	471.4
POLE-E14	5913.6	2696.5	9011.5	6653.9	3967.5	9931.1	675.5	73.6	1070.7	1149.4	196.8	1693.0	4962.1	2952.0	5940.0	4307.1	2299.5	5670.8	567.9	124.1	675.7	1451.6	745.0	1688.5
Total mean	2414.1	918.4	3578.8	2832.6	1412.9	4032.3	422.6	39.5	639.9	600.9	149.0	753.5	2759.8	1450.6	3645.0	2139.8	1107.1	2851.6	321.5	60.1	432.6	814.7	475.8	955.4
Percentage of exons with mean coverage >250 reads	98%			98%			71%			79%			97%			97%			54%			89%		
P25, 25 <sup>th</sup> percentile; P75, 75 <sup>th</sup> percentile. A mean coverage of >250 reads reflects a probability of 95% to detect a pathogenic variant [1]																								
[1] Eijkelenboom A, Kamping EJ, Kastner-van Raaij AW, Hendriks-Cornelissen SJ, Neveling K, Kuiper RP, et al. Reliable Next-Generation Sequencing of Formalin-Fixed, Paraffin-Embedded Tissue Using Single Molecule Tags. J Mol Diagn. 2016;18(6):851-63.																								

**Supplementary Table 3.** Baseline characteristics of each patient

Patient ID	Age	Histological diagnosis	Stage	Variants in surgical specimen	Variants in self-sample	Overlapping variants: surgical specimen and self-sample	Variants in Pap smear	Overlapping variants: surgical specimen and Pap smear	Variants in pipelle	Overlapping variants: surgical specimen and pipelle	Note
<b>Ovarian cancer patients</b>											
8	76	HGSC	3B	2	0	0	0	0	0	0	
10	41	MMMT	3B	1	0	0	0	0	0	0	
18	70	HGSC	3C	1	0	0	1	1	0	0	
20	65	HGSC	2A	1	0	0	1	1	0	0	
32	49	<i>Borderline</i>	<i>NA</i>	<i>0</i>	<i>0</i>	<i>0</i>	<i>0</i>	<i>0</i>	<i>0</i>	<i>0</i>	<i>Excluded*</i>
39	71	<i>Borderline serous</i>	<i>NA</i>	<i>0</i>	<i>0</i>	<i>0</i>	<i>0</i>	<i>0</i>	<i>0</i>	<i>0</i>	<i>Excluded*</i>
40	66	HGSC	3C	0	2	0	1	0	1	0	
44	32	HGSC	3C	0	0	0	1	0	0	0	
45	60	HGSC	2B	1	0	0	<i>NA</i>	<i>NA</i>	0	0	
49	50	Clear cell	3B	0	0	0	0	0	0	0	
54	61	<i>HGSC</i>	<i>4</i>	<i>0</i>	<i>0</i>	<i>0</i>	<i>0</i>	<i>0</i>	<i>0</i>	<i>0</i>	<i>Excluded*</i>
55	42	HGSC	3C	0	1	0	0	0	1	0	
58	55	Clear cell/serous	3A	0	1	0	1	0	0	0	
63	80	HGSC	1C	1	0	0	<i>NA</i>	<i>NA</i>	0	0	
65	58	HGSC	3C	1	0	0	0	0	<i>NA</i>	<i>NA</i>	
72	59	Clear cell/serous	4	1	0	0	0	0	0	0	
77	75	HGSC	3C	1	0	0	0	0	<i>NA</i>	<i>NA</i>	
85	59	HGSC	3C	1	0	0	0	0	0	0	
86	67	<i>LGSC</i>	<i>4</i>	<i>0</i>	<i>NA</i>	<i>NA</i>	<i>0</i>	<i>0</i>	<i>NA</i>	<i>NA</i>	<i>Excluded*</i>
107	71	HGSC	4	1	0	0	<i>NA</i>	<i>NA</i>	1	0	
114	77	<i>HGSC</i>	<i>99</i>	<i>0</i>	<i>1</i>	<i>0</i>	<i>NA</i>	<i>NA</i>	<i>0</i>	<i>0</i>	<i>Excluded*</i>
124	66	HGSC	3C	1	0	0	0	0	0	0	
127	54	Clear cell/ endometrioid	1C	2	0	0	3	1	<i>NA</i>	<i>NA</i>	

130	66	Adenosquamous	2B	2	0	0	0	0	NA	NA	
135	62	HGSC	3B	2	1	1	0	0	4	1	
151	69	HGSC	3C	1	0	0	2	1	0	0	
160	61	HGSC	3C	1	0	0	0	0	0	0	
164	58	HGSC	3C	0	0	0	0	0	NA	NA	Excluded*
165	74	Endometrioid	4	2	0	0	0	0	1	0	
167	69	Serous/ endometrioid	1A	3	0	0	0	0	1	0	
168	83	Endometrioid	2A	1	1	0	1	1	NA	NA	
181	76	HGSC	4	0	0	0	NA	NA	NA	NA	Excluded*
184	69	HGSC	3C	1	1	0	1	0	1	0	
188	68	HGSC	3C	1	0	0	0	0	0	0	
196	79	HGSC	3C	2	NA	NA	0	0	NA	NA	
197	66	HGSC	3C	0	NA	NA	0	0	NA	NA	Excluded*
198	65	HGSC	3C	3	NA	NA	3	3	0	0	
<b>Control patients</b>											
4	53	Myoma	-	0	0	0	0	0	0	0	
12	59	Mucinous cystadenoma	-	0	0	0	0	0	0	0	
28	57	Prolapse	-	0	0	0	0	0	0	0	
30	66	Simple cystadenoma	-	0	0	0	0	0	0	0	
33	71	Simple cystadenoma	-	0	0	0	0	0	0	0	
34	63	Teratoma	-	0	0	0	0	0	0	0	
38	53	Fibroma	-	0	0	0	1	0	0	0	
53	51	Teratoma	-	0	0	0	0	0	0	0	
60	57	Mucinous cystadenoma	-	0	0	0	0	0	1	0	
62	60	Fibroma	-	0	0	0	0	0	0	0	
67	45	Adenomyosis	-	0	0	0	0	0	0	0	

68	63	Mucinous cystadenoma	-	1	0	0	0	0	0	0	0
73	51	Myoma	-	0	0	0	0	0	0	0	0
74	72	Mucinous cystadenoma	-	1	0	0	0	0	0	0	0
76	59	Myoma	-	0	0	0	0	0	0	0	0
89	53	Teratoma	-	0	0	0	0	0	0	0	0
102	48	Myoma	-	0	0	0	0	0	0	0	0
104	53	Teratoma	-	0	0	0	0	0	0	0	0
115	63	Mucinous cystadenoma	-	0	0	0	0	0	0	0	0
116	58	Myoma	-	0	0	0	0	0	0	0	0
125	51	Mucinous cystadenoma	-	0	0	0	0	0	0	0	0
129	72	Serous cystadenoma	-	0	0	0	0	0	0	0	0
131	52	Myoma	-	0	0	0	0	0	0	0	0
133	62	Normal	-	0	0	0	0	0	0	0	0
136	48	Myoma	-	0	0	0	0	0	0	0	0
144	50	Myoma	-	0	0	0	0	0	0	0	0
150	46	Myoma	-	0	1	0	0	0	1	0	0
153	47	Normal	-	0	0	0	1	0	0	0	0
159	74	Prolapse	-	0	0	0	0	0	0	0	0
166	82	Inflammation	-	0	0	0	0	0	0	0	0
183	70	Fibroma	-	0	0	0	0	0	0	0	0
195		cystadenoma	-	NA	NA	NA	0	NA	0	NA	NA

\* Excluded from analysis as sequencing of the surgical specimen was unsuccessful: coverage was too low to potentially detect any variant  
 HGSC, high-grade serous cancer; LGSC, low-grade serous cancer; MMMT, Mixed Müllerian Tumor; NA, not applicable

**Supplementary Table 4. Identified pathogenic variants in the four specimens per patient**

Ovarian cancer patients						
Patient ID	Pathogenic variants	Pathogenic class	VAF in self-sample % (mutant reads)	VAF in Pap smear % (mutant reads)	VAF in pipelle % (mutant reads)	VAF in surgical specimen % (mutant reads)
8	PTEN:1008C>A p.(Tyr336*)	5	no	no	no	5.2 (38)
8	TP53:c.747G>T p.(Arg249Ser)	5	no	no	no	53.0 (184)
10	PTEN:c.389G>A p.(Arg130Gln)	5	no	no	no	5.7 (17)
18	TP53:c.743G>C p.(Arg248Pro)	5	no	0.37 (10)	no	52.0 (79)
20	TP53:c.743G>A p.(Arg248Gln)	4	no	0.46 (14)	no	100.0 (20)
32*	none		no	no	no	no
39*	none		no	no	no	no
40	PIK3CA:c.1633G>A p.(Glu545Lys)	5	1.3 (8)	no	no	no
40	TP53:c.*6T>C(3'UTR)		1.0(18)	5.4 (300)	8.8 (14)	no
44	TP53:c.574C>T p.(Gln192*)	5	no	1.4 (36)	no	no
45	TP53:c.814G>A p.(Val272Met)	5	no	NA	no	56.0 (379)
49	none		no	no	no	no
54*	none		no	no	no	no
55	KRAS:c.37G>T p.(Gly13Cys)	5	no	no	5.7 (97)	no
55	PIK3CA:c.1624G>A p.(Glu542Lys)	5	5.6 (32)	no	no	no
58	PIK3CA:c.1634A>G p.(Glu545Gly)	5	no	1.5 (19)	no	no
58	TP53:c.542G>A p.(Arg181His)	4	3.4 (18)	no	no	no
63	TP53:c.*6T>C(3'UTR)		no	NA	no	86.0 (111)
65	TP53:c.832C>G p.(Pro278Ala)	5	no	no	NA	73.0 (177)
72	TP53:c.711G>T p.(Met237Ile)	5	no	no	no	86.0 (89)
77	TP53:c.808_817del p.(Phe270fs)	5	no	no	NA	61.0 (81)
85	TP53:c.524G>A p.(Arg175His)	5	no	no	no	55.0 (12)
86*	none		NA	no	NA	no
107	PIK3CA:c.277C>T p.(Arg93Trp)	4	no	NA	no	13.0 (6)
107	TP53:c.818G>A p.(Arg273His)	5	no	NA	2.4 (14)	no
114*	ARID1A:c.2063A>G p.(His688Arg)	3	5.8 (12)	NA	no	no
124	TP53:c.672+1G>T p.?	5	no	no	no	55.0 (22)
127	ARID1A:c.2911G>A p.(Gly971Arg)	3	no	4.6 (42)	NA	no
127	PIK3CA:c.1634A>C p.(Glu545Ala)	5	no	0.24 (8)	NA	19.0 (235)
127	PIK3CA:c.1634A>G p.(Glu545Gly)	5	no	0.95 (32)	NA	no
127	TP53:c.427G>A p.(Val143Met)	5	no	no	NA	48.0 (68)
130	CTNBN1:c.121A>G p.(Thr41Ala)	5	no	no	NA	38.0 (1360)
130	PIK3CA:c.3203dup p.(Asn1068fs)	5	no	no	NA	36.0 (923)
135	ARID1A:c.4993+1G>A p.?	5	0.10 (8)	no	0.38 (6)	3.4 (6)
135	KRAS:c.37G>T p.(Gly13Cys)	5	no	no	35.0 (512)	no
135	PIK3CA:c.316G>C p.(Gly106Arg)	5	no	no	40.0 (323)	no
135	PTEN: c.968dup p.(Asn323fs)	5	no	no	80.0 (1117)	no
135	TP53:c.743G>A p.(Arg248Gln)	4	no	no	no	72.0 (186)
151	TP53:c.574C>T p.(Gln192*)	5	no	2.3 (33)	no	no

151	TP53:c.637_639delinsTGG p.(Arg213Trp)	5	no	0.18 (6)	no	90.0 (19)
160	TP53:c.743G>C p.(Arg248Pro)	5	no	no	no	71.0 (163)
164*	None		no	no	NA	no
165	CTNNB1:c.134C>T p.(Ser45Phe)	5	no	no	no	4.5 (26)
165	MTOR:c.4448G>A p.(Cys1483Tyr)	5	no	no	no	5.5 (62)
165	PIK3CA:c.1030G>A p.(Val344Met)	5	no	no	42.0 (1668)	no
167	CTNNB1:c.110C>G p.(Ser37Cys)	5	no	no	no	27.0 (499)
167	KRAS:c.35G>T p.(Gly12Val)	5	no	no	no	39.0 (1415)
167	PIK3CA:c.323G>A p.(Arg108His)	5	no	no	no	42.0 (1668)
167	PIK3CA:c.1625A>T p.(Glu542Val)	4	no	no	9.2 (40)	no
168	ARID1A:c.4101+2T>C p.?	5	1.2 (18)	no	NA	no
168	CTNNB1:c.110C>G p.(Ser37Cys)	5	no	0.10 (7)	NA	44.0 (72)
181*	none		no	NA	NA	no
184	PIK3CA:c.1634A>G p.(Glu545Gly)	5	0.4 (45)	0.17 (12)	4.2 (6)	no
184	TP53:c.839G>A p.(Arg280Lys)	4	no	no	no	41.0 (898)
188	TP53:c.724T>A p.(Cys242Ser)	4	no	no	no	87.0 (141)
196	PIK3CA:c.277C>T p.(Arg93Trp)	4	NA	no	NA	70.0 (1574)
196	TP53:c.747G>C p.(Arg249Ser)	5	NA	no	NA	62.0 (281)
197*	none		NA	no	NA	no
198	KRAS:c.35G>T p.(Gly12Val)	5	NA	4.7 (34)	no	13.0 (6)
198	PIK3CA:c.277C>T p.(Arg93Trp)	4	NA	6.15 (16)	no	27.0 (75)
198	PTEN:c.955_958del p.(Thr319*)	5	NA	12.0 (465)	no	63.0 (324)
<b>Control patients</b>						
4	none		no	no	no	no
12	none		no	no	no	no
28	none		no	no	no	no
30	none		no	no	no	no
33	none		no	no	no	no
34	none		no	no	no	no
38	PIK3CA:c.1634A>G p.(Glu545Gly)	5	no	1.8 (18)	no	no
53	none		no	no	no	no
60	PIK3CA:c.1634A>C p.(Glu545Ala)	5	no	no	3.7 (22)	no
62	none		no	no	no	no
67	none		no	no	no	no
68	KRAS:c.37G>T p.(Gly13Cys)	5	no	no	no	20.0 (303)
73	none		no	no	no	no
74	KRAS:c.35G>T p.(Gly12Val)	5	no	no	no	30.0 (44)
76	none		no	no	no	no
89	none		no	no	no	no
102	none		no	no	no	no
104	none		no	no	no	no
115	none		no	no	no	no
116	none		no	no	no	no
125	none		no	no	no	no

129	none		no	no	no	no
131	none		no	no	no	no
133	none		no	no	no	no
136	none		no	no	no	no
144	none		no	no	no	no
150	ARID1A:c.1270T>C p.(Ser424Pro)	3	2.5 (20)	no	no	no
150	PIK3CA:c.3203dup p.(Asn1068fs)	5	no	no	2.5 (10)	no
153	PIK3CA:c.1634A>G p.(Glu545Gly)	5	no	2.6 (6)	no	no
159	none		no	no	no	no
166	none		no	no	no	no
183	none		no	no	no	no
195	none		NA	no	no	NA
<p>* Excluded from analysis as sequencing of the surgical specimen was unsuccessful: coverage was too low to potentially detect any variant  VAF, variant allele frequency; NA, not applicable; Pathogenic class: 3, variant of unknown significance; 4, likely pathogenic; 5, pathogenic</p>						



**Supplementary Document 2. Detection rate per sampling method by stage**

In table:

	All stages		Early stage (I/II)		Late stage (III/IV)	
	n	%	n	%	n	%
Self-sample	6/27	22.2	1/7	14.3	5/20	25.0
Pap smear	10/26	38.5	3/5	60.0	7/21	33.3
Pipelle	7/23	30.4	1/4	25.0	6/19	31.6
Surgical specimen	24/29	82.8	7/7	100	17/22	77.3
Total	28/29	96.6	7/7	100	21/22	95.5

In figure:

