Association between the cervicovaginal microbiome, *BRCA1* mutation status, and risk of ovarian cancer: a case-control study

Nuno R Nené*, Daniel Reisel*, Andreas Leimbach*, Dorella Franchi, Allison Jones, Iona Evans, Susanne Knapp, Andy Ryan, Shohreh Ghazali, John F Timms, Tobias Paprotka, Line Bjørge, Michal Zikan, David Cibula, Nicoletta Colombo, Martin Widschwendter

Summary

Background Various factors—including age, family history, inflammation, reproductive factors, and tubal ligation—modulate the risk of ovarian cancer. In this study, our aim was to establish whether women with, or at risk of developing, ovarian cancer have an imbalanced cervicovaginal microbiome.

Methods We did a case-control study in two sets of women aged 18–87 years in the Czech Republic, Germany, Italy, Norway, and the UK. The ovarian cancer set comprised women with epithelial ovarian cancer and controls (both healthy controls and those diagnosed with benign gynaecological conditions). The BRCA set comprised women with a *BRCA1* mutation but without ovarian cancer and controls who were wild type for *BRCA1* and *BRCA2* (both healthy controls and those with benign gynaecological conditions). Cervicovaginal samples were gathered from all participants with the ThinPrep system and then underwent 16S rRNA gene sequencing. For each sample, we calculated the proportion of lactobacilli species (ie, *Lactobacillus crispatus*, *Lactobacillus iners*, *Lactobacillus gasseri*, and *Lactobacillus jensenii*), which are essential for the generation of a protective low vaginal pH, in the cervicovaginal microbiota. We grouped samples into those in which lactobacilli accounted for at least 50% of the species present (community type L) and those in which lactobacilli accounted for less than 50% of the species present (community type O). We assessed the adjusted association between *BRCA1* status and ovarian cancer status and cervicovaginal microbiota community type, using a logistic regression model with a bias reduction method.

Findings Participants were recruited between Jan 2, 2016, and July 21, 2018. The ovarian cancer set (n=360) comprised 176 women with epithelial ovarian cancer, 115 healthy controls and 69 controls with benign gynaecological conditions. The BRCA set (n=220) included 109 women with *BRCA1* mutations, 97 healthy controls wild type for *BRCA1* and *BRCA2* and 14 controls with a benign gynaecological condition wild type for *BRCA1* and *BRCA2*. On the basis of two-dimensional density plots, receiver–operating characteristic curve analysis, and age thresholds used previously, we divided the cohort into those younger than 50 years and those aged 50 years or older. In the ovarian cancer set, women aged 50 years or older had a higher prevalence of community type O microbiota (81 [61%] of 133 ovarian cancer cases and 84 [59%] of 142 healthy controls) than those younger than 50 years (23 [53%] of 43 cases and 12 [29%] of 42 controls). In the ovarian cancer set, women younger than 50 years with ovarian cancer had a significantly higher prevalence of community type O microbiota than did age-matched controls under a logistic regression model with bias correction (odds ratio [OR] 2·80 [95% CI 1·17–6·94]; p=0·020). In the BRCA set, women with *BRCA1* mutations younger than 50 years were also more likely to have community type O microbiota than age-matched controls (OR 2·79 [95% CI 1·25–6·68]; p=0·012), after adjustment for pregnancy (ever). This risk was increased further if more than one first-degree family member was affected by any cancer (OR 5·26 [95% CI 1·83–15·30]; p=0·0022). In both sets, we noted that the younger the participants, the stronger the association between community type O microbiota and ovarian cancer or *BRCA1* mutation status (eg, OR for community type O for cases aged <40 years in the ovarian cancer set 7·00 [95% CI 1·27–51·44], p=0·025; OR for community type O for *BRCA1* mutation carriers aged <35 years in the BRCA set 4·40 [1·14–24·36], p=0·031).

Interpretation The presence of ovarian cancer, or factors known to affect risk for the disease (ie, age and *BRCA1* germline mutations), were significantly associated with having a community type O cervicovaginal microbiota. Whether re-instatement of a community type L microbiome by using, for example, vaginal suppositories containing live lactobacilli, would alter the microbiomial composition higher up in the female genital tract and in the fallopian tubes (the site of origin of high-grade serous ovarian cancer), and whether such changes could translate into a reduced incidence of ovarian cancer, needs to be investigated.

Funding EU Horizon 2020 Research and Innovation Programme, EU Horizon 2020 European Research Council Programme, and The Eve Appeal.

Department of Women's Cancer, EGA Institute for Women's Health, Faculty of Population Health Sciences, University College London, London, UK (N R Nené PhD, D Reisel PhD, A Jones BSc, I Evans PhD, S Knapp PhD, A Ryan PhD, S Ghazali MSc, J F Timms DPhil, Prof M Widschwendter MD); **Eurofins Genomics Europe Sequencing, Constance, Germany** (A Leimbach PhD, T Paprotka PhD); **Istituto Europeo di Oncologia, IRCCS, Milan, Italy** (D Franchi MD, Prof N Colombo MD); **Department of Obstetrics and Gynaecology, Haukeland University Hospital, Bergen, Norway** (Prof L Bjørge PhD); **Centre for Cancer Biomarkers, Department of Clinical Science, University of Bergen, Bergen, Norway** (Prof L Bjørge); **Gynaecologic Oncology Center** (Prof D Cibula MD) **and Department of Gynecology and Obstetrics** (Prof M Zikan MD), **First Faculty of Medicine, Charles University in Prague, Prague, Czech Republic; Hospital Na Bulovce, Prague, Czech Republic** (Prof M Zikan); **General University Hospital in Prague, Prague, Czech Republic** (Prof D Cibula); **and University of Milano-Bicocca, Milan, Italy** (Prof N Colombo)

Correspondence to:
Prof Martin Widschwendter,
Department of Women's Cancer,
EGA Institute for Women's Health,
Faculty of Population Health
Sciences, University College
London, London WC1E 6AU, UK
m.widschwendter@ucl.ac.uk

Research in context

Evidence before this study

At the initiation of this case-control study, on Sept 15, 2015, we searched PubMed with the terms "[ovarian cancer or BRCA1] and [cervical or vaginal] and [microbiome or lactobacilli]" for studies published in any language in which the vaginal or cervical microbiome was assessed in women with ovarian cancer or *BRCA1* mutations. We also manually searched the reference lists of papers identified by our search. Neither the original search, nor any subsequent searches, the last of which was done on Dec 20, 2018, identified any published studies about this topic.

Added value of this study

To our knowledge, this study is the first case-control analysis of the cervicovaginal microbiome in women with ovarian cancer and in women at high risk of ovarian cancer (ie, with a *BRCA1* mutation). Our finding that a dysbiosis (ie, a lower proportion of lactobacilli in the cervicovaginal microbiome) is present not

only in women younger than 50 years with established ovarian cancer, but also in those at high risk of the disease, supports the view that various factors, including genetic factors, reveal their carcinogenic potential via a modifiable cell non-autonomous mechanism, such as the microbiome.

Implications of all the available evidence

Although our findings suggest that the cervicovaginal microbiome is implicated in ovarian cancer risk, the causative link between dysbiosis and ovarian cancer has yet to be further assessed. The proportion of lactobacilli should be combined with genetic and epidemiological factors in ovarian cancer risk predicting models. Once the functional relevance of community type O cervicovaginal microbiota in ovarian cancer development has been established, vaginal lactobacilli transplantation in young carriers of germline *BRCA1* mutations could become available as a risk-reduction measure.

Introduction

Evidence is accumulating that microbiota can affect the risk of cancer development, response to oncological therapy, and development of cancer-associated and treatment-related complications.[1] Most of the available evidence so far relates to the gut microbiome,[2] but some data suggest roles for the urinary[3] and cervicovaginal[4] microbiomes in the development of bladder and endometrial cancers, respectively. Whether the absence or disruption of the normal protective microbiome—and particularly of the non-gut microbiome—is associated with predisposition to cancer has not been thoroughly investigated. This information is a prerequisite for assessment of the risk associated with dysbiosis and to assess whether restoration of the normal microbiota could help to prevent cancer.

In addition to age, *BRCA1* and *BRCA2* germline mutations, parity, oral contraceptive use, and endometriosis,[5] ascending infections leading to pelvic inflammatory disease[6] and procedures that theoretically mitigate the risk of ascending infection (ie, tubal ligation) modify the risk of ovarian cancer. The cervicovaginal microbiota is dominated by *Lactobacillus crispatus*, *Lactobacillus iners*, *Lactobacillus gasseri*, or *Lactobacillus jensenii*,[7–9] and prevents urogenital disease by lowering the vaginal pH via lactic acid production and by several other mechanisms, including immunomodulation.[10] Ravel and colleagues[7] described four types of *Lactobacillus* community (groups I, II, III, and V, which are dominated by *L crispatus*, *L gasseri*, *L iners*, and *L jensenii*, respectively) that—compared with community type IV, which contains higher proportions of strictly anaerobic organisms—are associated with a substantially lower vaginal pH, which could potentially decrease the risk of ascending infection.

In this case-control study, we aimed to assess whether the cervicovaginal microbiota was altered in women who had recently received a diagnosis of ovarian cancer, and whether alterations were already present in women with a germline *BRCA1* mutation, who have a 40-times increased risk of ovarian cancer,[11] who had not yet developed ovarian cancer.

Methods

Study design and participants

We did a case-control study of the association between cervicovaginal microbiota composition, *BRCA1* mutation status, and ovarian cancer status, in two separate datasets, by using 16S rRNA sequencing. The ovarian cancer set comprised women with epithelial ovarian cancer and controls (both healthy controls and those diagnosed with benign gynaecological conditions). The BRCA set comprised women with a *BRCA1* mutation but without ovarian cancer and controls who were wild type for *BRCA1* (both healthy controls and those with benign gynaecological conditions). Our study was part of a multicentre study (the FORECEE [4C] Programme), which was done at 14 recruitment sites in five European countries: the Czech Republic, Germany, Italy, Norway, and the UK (appendix pp 5–6). Participants aged 18–87 years were recruited to our microbiome substudy in three different settings: women with suspected ovarian cancer, benign gynaecological conditions, and women with *BRCA1* mutations were recruited from hospitals during outpatient clinics; healthy women from the general population were recruited via routine cervical cancer screening programmes; and both women who were *BRCA1* mutation carriers and healthy women were recruited by a dedicated research clinic (through outreach campaigns and public engagement). Eligible ovarian cancer cases were women with a suspected diagnosis of a

See **Online** for appendix

malignant invasive epithelial ovarian cancer who were recruited before undergoing surgery or receiving any chemotherapy or radiotherapy. Before agreeing to take part, each prospective participant was given a participant information sheet, and the rationale for the study was explained. Additional resources, including an explanatory video and further online resources, were also made available. Before enrolment, prospective participants completed an epidemiological questionnaire (administered via the Qualtrics application on dedicated iPads), comprising questions about current and historical health habits, relevant risk factors, and their medical and obstetric history. All participants provided written informed consent. The FORECEE (4C) Programme received ethical approval from the UK Health Research Authority (REC 14/LO/1633).

Procedures

A small group of research midwives or physicians used the ThinPrep system (Hologic, Marlborough, MA, USA) to collect cervical smear samples from all participants at collaborating hospitals and recruitment centres. Briefly, cervical cells were sampled with a Rovers cervix brush (Rovers Medical Devices, Oss, Netherlands), which was inserted into the vagina, rotated fully five times while in contact with the cervix, then removed and immersed in a ThinPrep vial containing PreservCyt fluid (Hologic). The brush was pushed against the bottom of the vial ten times to facilitate release of the cells into the PreservCyt solution, and then discarded. Sample vials were sealed and stored locally at room temperature before shipment to University College London (London, UK) at ambient temperature. All biological samples were assigned anonymous participant identification numbers, which were matched with participants' names in a securely stored link file.

During wet laboratory processing, each cervical smear sample was poured into 50 mL Falcon tubes (Fisher Scientific, Loughborough, UK) and left to sediment at room temperature for 2 h. 1 mL wide-bore tips were used to transfer cervical cell enriched sediment to 2 mL cryovials, which were stored at –80°C. Next-generation sequencing was done on all samples in both sets. Total DNA extraction from cervical smears was done with the QIAsymphony DSP Virus/Pathogen kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. Sequencing and taxonomical classification of bacterial species in cervical samples was done by Eurofins Genomics Europe Sequencing (Constance, Germany) as previously described.[12] In brief, the hypervariable V1–V3 region of the bacterial small subunit ribosomal DNA (16S rRNA genes) was amplified with primers modified from work by Klindworth and colleagues (27F 5′-AGAGTTTGATCCTGGCTCAG and 534R 5′-ATTACCGCGGCTGCTGG) to decrease melting temperature stringency and, as a result, increase taxonomic coverage, including Illumina adapters and dual index barcodes. These barcoded libraries were equimolarly pooled and sequenced on the HiSeq2500 with 300 bp paired-end reads (HiSeq Rapid SBS Kit v2; Illumina, San Diego, CA, USA). The 16S rRNA gene PCRs, next-generation sequencing libraries, and pooled libraries were quality checked by Eurofins Genomics via electropherograms and fluorometer concentration determinations. Additionally, positive and negative controls were included in each library preparation, including DNA extraction batches.

To process the 16S rRNA gene-sequencing data, the demultiplexed sequencing reads were quality checked, trimmed, and filtered with Sickle (version 1.33; settings: -q 20, -l 246, -n yes), and adapters and primers were removed with Cutadapt (version 1.10; settings: --minimum-length 246). Overlapping paired-end reads were merged for full-length V1–V3 16S amplicons with FLASH (version 1.2.11; settings: --min-overlap 10, --max-overlap 300, --max-mismatch-density 0·25)[14] and clustered with CD-HIT (version 4.6; settings: -c 0·99 and cluster filtered, resulting in a minimum number of two members),[15] and then chimeric sequences were removed with UCHIME (version 4.2.40).[16] Operational taxonomic units were assigned with BLASTN+ (version 2.4.0; -evalue 1e-06)[17] via a non-redundant 16S rRNA reference database from the Ribosomal Database Project (release 11),[18] and filtered to ensure that only high-quality operational taxonomic units were included (ie, 97% identity threshold, 95% alignment coverage, 376 query length threshold, and 10% bitscore threshold). Taxonomic classification was based on the US National Center for Biotechnology Information taxonomy.[19] To establish species diversity and evenness, we calculated both the Shannon and the Simpson indices for each sample.[7]

For the purposes of this study, we collapsed Ravel's four *Lactobacillus* community groups I, II, III, and V into one microbial community, which we referred to as community type L. We divided samples into people whose cervicovaginal microbiota consisted of at least 50% community type L and those whose microbiota consisted of less than 50% community type L, which we referred to as community type O. Generally, community type L microbiota are dominated by a *Lactobacillus* species, whereas type O microbiota communities have more bacterial diversity, are composed of typical obligate and facultative anaerobe genera (such as *Gardnerella* or *Atopobium* species), and are in turn associated with aerobic vaginitis and bacterial vaginosis.[7]

For control participants in the BRCA set, next-generation sequencing of white blood cell DNA (blood samples were taken from all participants in the FORECEE [4C] Programme) was done by Color Genomics (Burlingame, CA, USA) to exclude a germline mutation in *BRCA1* and *BRCA2*. Mutation testing was not done for control participants in the ovarian cancer set.

Statistical analysis

As part of the FORECEE programme, we estimated that we required 300–400 cases and controls for the ovarian cancer set. The power calculation used to determine the proposed sample sizes was based on methods used to quantify sample size for analyses of DNA methylation.[20]

	Age <50 years			Age ≥50 years		
	Controls (n=42)	Cases (n=43)	p value	Controls (n=142)	Cases (n=133)	p value
Ethnicity	..	..	0·21	..	..	0·061
White	40 (95%)	43 (100%)	..	126 (89%)	127 (95%)	..
Non-white	2 (5%)	0	..	15 (11%)	6 (5%)	..
Missing	0	0	..	1 (1%)	0	..
Body-mass index (kg/m^2)	..	..	0·58	..	..	0·31
≤30	37 (88%)	36 (84%)	..	111 (78%)	97 (73%)	..
>30	5 (12%)	7 (16%)	..	31 (22%)	36 (27%)	..
Current smoker	..	..	0·10	..	..	0·029
No	37 (88%)	31 (72%)	..	125 (88%)	110 (83%)	..
Yes	5 (12%)	11 (26%)	..	10 (7%)	21 (16%)	..
Missing	0	1 (2%)	..	7 (5%)	2 (2%)	..
Age at menarche (years)	..	..	0·23	..	..	0·32
≤12	15 (36%)	21 (49%)	..	63 (44%)	52 (39%)	..
>12	27 (64%)	22 (51%)	..	76 (54%)	80 (60%)	..
Missing	0	0	..	3 (2%)	1 (1%)	..
Duration of OCP use (years)	..	..	0·038	..	..	0·038
<5	21 (50%)	34 (79%)	..	101 (71%)	111 (83%)	..
≥5	18 (43%)	8 (19%)	..	32 (23%)	18 (14%)	..
Missing	3 (7%)	1 (2%)	..	9 (6%)	4 (3%)	..
Current OCP use	..	..	0·039	..	..	0·18
No	32 (76%)	41 (95%)	..	123 (87%)	125 (94%)	..
Yes	6 (14%)	1 (2%)	..	9 (6%)	4 (3%)	..
Missing	4 (10%)	1 (2%)	..	10 (7%)	4 (3%)	..
Ever pregnant	..	..	0·81	..	..	0·039
No	12 (29%)	13 (30%)	..	11 (8%)	21 (16%)	..
Yes	30 (71%)	29 (67%)	..	130 (92%)	111 (83%)	..
Missing	0	1 (2%)	..	1 (1%)	1 (1%)	..
Menopausal status	..	..	0·96	..	..	0·42
Premenopausal	34 (81%)	35 (81%)	..	11 (8%)	7 (5%)	..
Postmenopausal	8 (19%)	8 (19%)	..	131 (92%)	126 (95%)	..
Ever used HRT	..	..	0·021	..	..	0·0080
No	37 (88%)	43 (100%)	..	97 (68%)	109 (82%)	..
Yes	5 (12%)	0	..	44 (31%)	23 (17%)	..
Missing	0	0	..	1 (1%)	1 (1%)	..
Current combined hormone use	..	..	0·039	..	..	0·46
No	32 (76%)	41 (95%)	..	128 (90%)	127 (95%)	..
Yes	6 (14%)	1 (2%)	..	4 (3%)	2 (2%)	..
Missing	4 (10%)	1 (2%)	..	10 (7%)	4 (3%)	..
Duration of combined hormone use (years)	..	..	0·011	..	..	0·0020
<5	22 (52%)	33 (77%)	..	97 (68%)	113 (85%)	..
≥5	17 (40%)	9 (21%)	..	34 (24%)	14 (11%)	..
Missing	3 (7%)	1 (2%)	..	11 (8%)	6 (5%)	..
						(Table 1 continues on next page)

	Age <50 years			Age ≥50 years		
	Controls (n=42)	Cases (n=43)	p value	Controls (n=142)	Cases (n=133)	p value
(Continued from previous page)						
Microbiota community type	..	..	0·020	..	..	0·77
L	30 (71%)	20 (47%)	..	58 (41%)	52 (39%)	..
O	12 (29%)	23 (53%)	..	84 (59%)	81 (61%)	..

We calculated p values under a logistic regression model with a bias reduction method. The appendix shows baseline characteristics for hospital-based controls with benign gynaecological conditions (p 7) and missing values per covariate (p 23). OCP=oral contraceptive pill. HRT=hormone replacement therapy.

Table 1: Descriptive characteristics of cases and controls in the ovarian cancer set for all covariates used in prediction models

Because the availability of microbiome data for each participant was part of a concerted effort to establish a multiomics pipeline for women's cancer prevention, the same operating characteristics used in the other parts of FORECEE were assumed for the datasets reported here.

Controls were initially matched one to one with cases (ovarian cancer cases or *BRCA1* mutation carriers) based on menopausal status, age (in 5-year age ranges when possible), and recruitment centre and country.

To classify and identify predictors, we fitted a logistic regression model independently to each covariate in the clinical data collected from each participant in the study, to *BRCA1*, ovarian cancer, or control status, and to microbiota community status (ie, type L or type O). The covariates leading to a significant association with the outcome (ie, a $p<0·05$, with the 95% CI for the calculcated odds ratio [OR] not crossing 1) were selected to adjust the prediction performance of *BRCA1*, ovarian cancer status, and microbiota community type status. We tested both additive models with and without interaction terms. The logistic regression model was implemented in the *logistf* R package (version 1.23), which fits a logistic regression model according to Firth's bias reduction method. Reported 95% CIs and tests were based on the profile penalised log likelihood. Given that the number of samples was larger than the number of predictors, even when adjusted (maximum of six covariates selected), the logistic regression model was stable. Cases with missing values were omitted in each independent fitting, but not across the whole study. Forest plots were created with the *forestplot* R package (version 1.9).

Receiver-operator characteristic curve analysis and the Youden's J statistic were used to calculate the optimal age threshold by which to divide the cohorts into separate age groups. Abundance patterns within each age group and each microbiota community type were clustered by a hierarchical clustering algorithm, *hclust*, in R, via the Ward's method. The patterns were scaled column-wise. The species selected for the heatmaps correspond to *L crispatus, L iners, L gasseri, or L jensenii* (ie, community type L), in addition to those that ranked highest in terms of average abundance across all participants in each set.

To assess whether sample transit times and time to pre-processing affected our results, we did a two-tailed t test (in SPSS, version 25.0.0.1) for the equality of means, in which we compared time from taking of samples to receipt of the samples by University College London, time from taking samples to pre-processing, and time between taking samples and final analysis, between cases and controls, in both datasets. A univariate logistic regression test with either cancer status or mutation status as the response variable and each of the time variables as the sole predictors was also used to confirm the significance of the results. Additionally, we also tested under the logistic regression paradigm the predictive capacity of the these times when the response variable was microbiota community type.

Role of the funding source

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results

Between Jan 2, 2016, and May 28, 2018, we recruited 219 cases with ovarian cancer and 219 age-matched controls, who comprised the ovarian cancer set. Of these participants, 176 cases with epithelial ovarian cancer and 184 controls (115 healthy controls from the general population and 69 controls with benign gynaecological conditions; table 1; appendix p 7) provided sufficient DNA for 16S rRNA gene sequencing. Of the 176 women with ovarian cancer, 119 (68%) had serous high-grade disease, 15 (9%) had clear cell disease, 16 (9%) had endometrioid disease, 13 (7%) had mucinous disease, and 13 (7%) had serous low-grade disease. 26 women (15%) had grade 1 ovarian cancer, 14 (8%) had grade 2 disease, and 131 (74%) had grade 3; five (3%) were not graded. 66 (38%) had stage I or II disease and 108 (61%) had stage III or IV disease; two (1%) were not staged.

Between Feb 7, 2016, and July 21, 2018, we recruited 131 cases with a *BRCA1* mutation and 131 age-matched controls, who together comprised the BRCA set. We analysed the microbiota of 109 *BRCA1* mutation carriers and 111 age-matched controls (97 healthy controls who were wild type for *BRCA1,* and 14 controls with a benign

	Age <50 years			Age ≥50 years		
	Controls (n=86)	Cases (n=94)	p value	Controls (n=25)	Cases (n=15)	p value
Ethnicity	..	..	0·40	..	..	0·72
White	80 (93%)	84 (89%)	..	21 (84%)	12 (80%)	..
Non-white	6 (7%)	10 (11%)	..	4 (16%)	3 (20%)	..
BMI (kg/m²)	..	..	0·19	..	..	0·56
≤30	78 (91%)	79 (84%)	..	22 (88%)	12 (80%)	..
>30	8 (9%)	15 (16%)	..	3 (12%)	3 (20%)	..
Current smoker	..	..	0·83	..	..	0·51
No	77 (90%)	86 (91%)	..	23 (92%)	12 (80%)	..
Yes	8 (9%)	8 (9%)	..	2 (8%)	2 (13%)	..
Missing	1 (1%)	0	..	0	2 (13%)	..
Age at menarche (years)	..	..	0·69	..	..	0·93
≤12	40 (47%)	40 (43%)	..	10 (40%)	6 (40%)	..
>12	46 (53%)	52 (55%)	..	14 (55%)	9 (60%)	..
Missing	0	2 (2%)	..	1 (4%)	0	..
Duration of OCP use (years)	..	..	0·32	..	..	0·48
<5	37 (43%)	35 (37%)	..	15 (60%)	6 (40%)	..
≥5	41 (48%)	53 (56%)	..	9 (36%)	6 (40%)	..
Missing	8 (9%)	6 (6%)	..	1 (4%)	3 (20%)	..
Current OCP use	..	..	0·50	..	..	0·11
No	57 (66%)	60 (64%)	..	19 (76%)	12 (80%)	..
Yes	21 (24%)	28 (30%)	..	5 (20%)	0	..
Missing	8 (9%)	6 (6%)	..	1 (4%)	3 (20%)	..
Ever pregnant	..	..	0·51	..	..	0·40
No	39 (45%)	38 (40%)	..	2 (8%)	0	..
Yes	47 (55%)	56 (60%)	..	23 (92%)	15 (100%)	..
Menopausal status	..	..	0·016	..	..	0·82
Premenopausal	86 (100%)	88 (94%)	..	6 (24%)	3 (20%)	..
Postmenopausal	0	6 (6%)	..	19 (76%)	12 (80%)	..
Ever used HRT	..	..	0·92	..	..	0·82
No	84 (98%)	92 (98%)	..	19 (76%)	11 (73%)	..
Yes	2 (2%)	2 (2%)	..	6 (24%)	4 (27%)	..
Current combined hormone use	..	..	0·94	..	..	0·13
No	69 (80%)	75 (80%)	..	21 (84%)	15 (100%)	..
Yes	17 (20%)	19 (20%)	..	4 (16%)	0	..
Duration of combined hormone use (years)	..	..	0·69	..	..	0·91
<5	51 (59%)	53 (56%)	..	17 (68%)	10 (67%)	..
≥5	35 (41%)	41 (44%)	..	8 (32%)	5 (33%)	..
Microbiota community type	..	..	0·0090	..	..	0·99
L	77 (90%)	70 (74%)	..	10 (40%)	6 (40%)	..
O	9 (10%)	24 (26%)	..	15 (60%)	9 (60%)	..

We calculated p values under a logistic regression model with a bias reduction method. The appendix shows baseline characteristics for hospital-based controls with benign gynaecological conditions (p 7) and missing values per covariate (p 23). OCP=oral contraceptive pill. HRT=hormone replacement therapy.

Table 2: Descriptive characteristics of cases and controls in the BRCA set for all covariates used in prediction models

gynaecological condition wild type for *BRCA1*) who provided sufficient amounts of DNA for the microbiome analyses (table 2; appendix p 7).

Although we had planned to match cases and controls in both the ovarian cancer set and the BRCA set on the basis of menopausal status, age, and recruitment centre
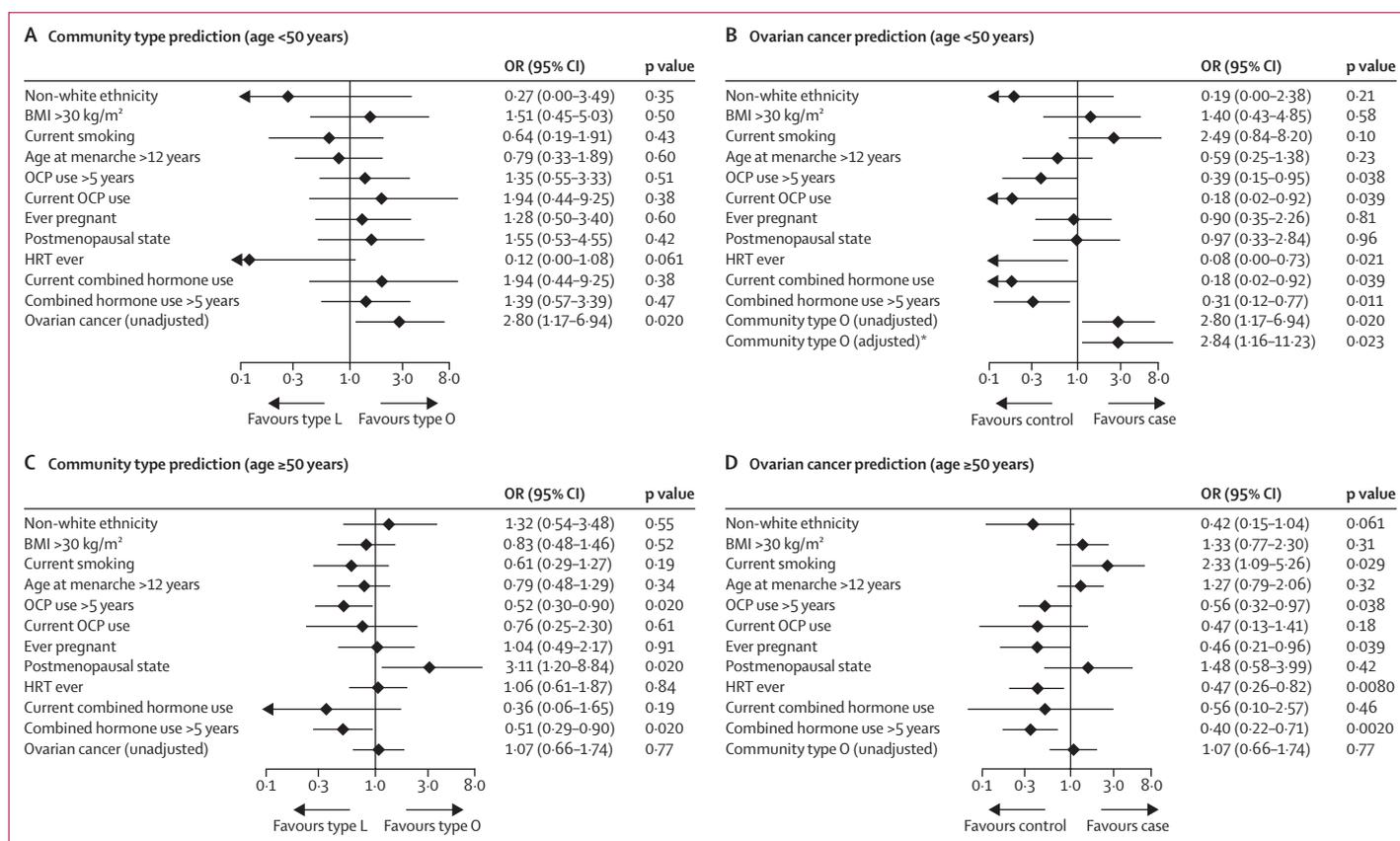
A Community type prediction (age <50 years)

	OR (95% CI)	p value
Non-white ethnicity	0·27 (0·00–3·49)	0·35
BMI >30 kg/m²	1·51 (0·45–5·03)	0·50
Current smoking	0·64 (0·19–1·91)	0·43
Age at menarche >12 years	0·79 (0·33–1·89)	0·60
OCP use >5 years	1·35 (0·55–3·33)	0·51
Current OCP use	1·94 (0·44–9·25)	0·38
Ever pregnant	1·28 (0·50–3·40)	0·60
Postmenopausal state	1·55 (0·53–4·55)	0·42
HRT ever	0·12 (0·00–1·08)	0·061
Current combined hormone use	1·94 (0·44–9·25)	0·38
Combined hormone use >5 years	1·39 (0·57–3·39)	0·47
Ovarian cancer (unadjusted)	2·80 (1·17–6·94)	0·020

Favours type L Favours type O

B Ovarian cancer prediction (age <50 years)

	OR (95% CI)	p value
Non-white ethnicity	0·19 (0·00–2·38)	0·21
BMI >30 kg/m²	1·40 (0·43–4·85)	0·58
Current smoking	2·49 (0·84–8·20)	0·10
Age at menarche >12 years	0·59 (0·25–1·38)	0·23
OCP use >5 years	0·39 (0·15–0·95)	0·038
Current OCP use	0·18 (0·02–0·92)	0·039
Ever pregnant	0·90 (0·35–2·26)	0·81
Postmenopausal state	0·97 (0·33–2·84)	0·96
HRT ever	0·08 (0·00–0·73)	0·021
Current combined hormone use	0·18 (0·02–0·92)	0·039
Combined hormone use >5 years	0·31 (0·12–0·77)	0·011
Community type O (unadjusted)	2·80 (1·17–6·94)	0·020
Community type O (adjusted)*	2·84 (1·16–11·23)	0·023

Favours control Favours case

C Community type prediction (age ≥50 years)

	OR (95% CI)	p value
Non-white ethnicity	1·32 (0·54–3·48)	0·55
BMI >30 kg/m²	0·83 (0·48–1·46)	0·52
Current smoking	0·61 (0·29–1·27)	0·19
Age at menarche >12 years	0·79 (0·48–1·29)	0·34
OCP use >5 years	0·52 (0·30–0·90)	0·020
Current OCP use	0·76 (0·25–2·30)	0·61
Ever pregnant	1·04 (0·49–2·17)	0·91
Postmenopausal state	3·11 (1·20–8·84)	0·020
HRT ever	1·06 (0·61–1·87)	0·84
Current combined hormone use	0·36 (0·06–1·65)	0·19
Combined hormone use >5 years	0·51 (0·29–0·90)	0·020
Ovarian cancer (unadjusted)	1·07 (0·66–1·74)	0·77

Favours type L Favours type O

D Ovarian cancer prediction (age ≥50 years)

	OR (95% CI)	p value
Non-white ethnicity	0·42 (0·15–1·04)	0·061
BMI >30 kg/m²	1·33 (0·77–2·30)	0·31
Current smoking	2·33 (1·09–5·26)	0·029
Age at menarche >12 years	1·27 (0·79–2·06)	0·32
OCP use >5 years	0·56 (0·32–0·97)	0·038
Current OCP use	0·47 (0·13–1·41)	0·18
Ever pregnant	0·46 (0·21–0·96)	0·039
Postmenopausal state	1·48 (0·58–3·99)	0·42
HRT ever	0·47 (0·26–0·82)	0·0080
Current combined hormone use	0·56 (0·10–2·57)	0·46
Combined hormone use >5 years	0·40 (0·22–0·71)	0·0020
Community type O (unadjusted)	1·07 (0·66–1·74)	0·77

Favours control Favours case

Figure 1: Effect of selected covariates on microbiota community type and risk of ovarian cancer in the ovarian cancer set
(A) Prediction of microbiota community type in participants younger than 50 years. (B) Prediction of ovarian cancer in participants younger than 50 years. (C) Prediction of microbiota community type in participants aged 50 years or older. (D) Prediction of ovarian cancer in participants aged 50 years or older. Arrows represent truncated 95% CIs. More data are presented in the appendix (pp 8–12). OR=odds ratio. BMI=body-mass index. OCP=oral contraceptive pill. HRT=hormone replacement therapy. *Adjusted for current OCP use, duration of OCP, HRT use (ever), current combined hormone use, and duration of combined hormone use.

and country, because of an imbalance in recruitment of cases and controls at some centres, some cases were matched according to age and menopausal status alone (appendix pp 23–24).

On the basis of two-dimensional density plots (appendix p 2), receiver–operating characteristic curve analysis, and age thresholds used previously,[7] we divided the cohorts into two age groups: those younger than 50 years and those aged 50 years or older. The choice of the threshold of 50 years was a compromise between 45 years—the upper limit for the recruitment of volunteers reported previously[7]—and 55·2 years, which was the value determined by fitting a univariate logistic regression model to the controls in the ovarian cancer set with microbiota community type as the response variable and age as the sole predictor. The two-dimensional densities (appendix p 2), for which the sum of percentages of species associated with community type L was plotted against age for all participants overall and in the ovarian cancer and BRCA sets individually, showed that the threshold of 50 years distinguished the peaks associated with very low microbiota community type L percentages.

Of the 360 people in the ovarian cancer set in whom 16S rRNA gene sequencing was done, 275 were aged 50 year or older (133 cases and 142 controls) and 85 were younger than 50 years (43 cases and 42 controls; table 1). 165 (60%) of the 275 women aged 50 years or older had community type O microbiota, compared with 35 (41%) of the 85 women younger than 50 years (table 1). 26 (39%) of 66 patients with stage I or II ovarian cancer, and 45 (42%) of 108 with stage III or IV ovarian cancer, had community type L microbiota.

In the ovarian cancer set, the mean time from taking of smear samples to delivery of samples to University College London for testing was 13·5 days (SD 11·5; range 0–71), and the mean time from receipt of samples to pre-processing (ie, enrichment and storage at –80°C) was 23·5 days (SD 34·0; range 0–434 days). A *t* test showed that neither time from collection to receipt (p=0·66) nor time from receipt to pre-processing (p=0·12) differed significantly between cases and controls.

In the BRCA set, the mean time from collection of samples to receipt was 10·8 days (SD 9·2; range 0–38 days), and from receipt of samples to pre-processing
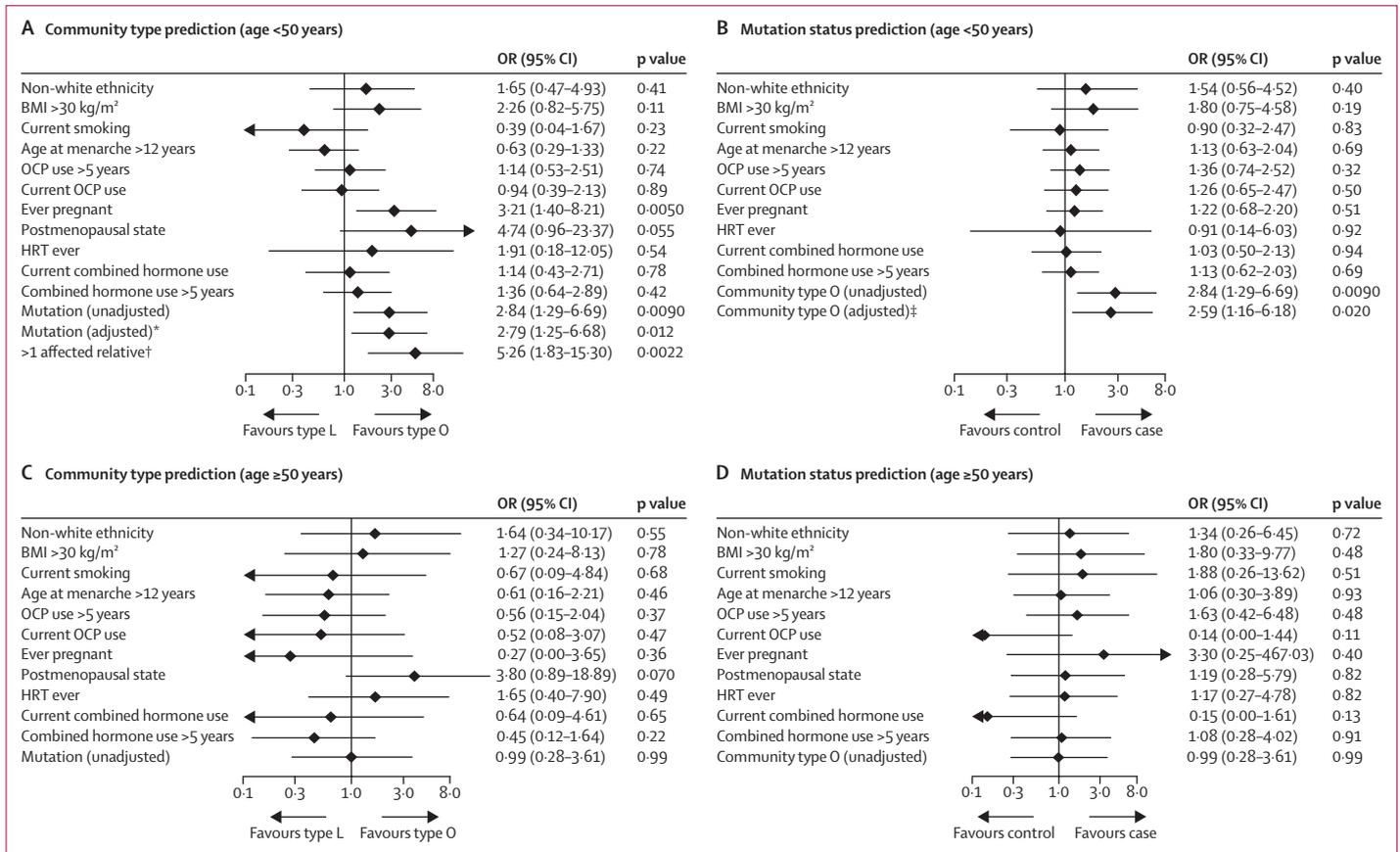
A Community type prediction (age <50 years)

	OR (95% CI)	p value
Non-white ethnicity	1·65 (0·47–4·93)	0·41
BMI >30 kg/m²	2·26 (0·82–5·75)	0·11
Current smoking	0·39 (0·04–1·67)	0·23
Age at menarche >12 years	0·63 (0·29–1·33)	0·22
OCP use >5 years	1·14 (0·53–2·51)	0·74
Current OCP use	0·94 (0·39–2·13)	0·89
Ever pregnant	3·21 (1·40–8·21)	0·0050
Postmenopausal state	4·74 (0·96–23·37)	0·055
HRT ever	1·91 (0·18–12·05)	0·54
Current combined hormone use	1·14 (0·43–2·71)	0·78
Combined hormone use >5 years	1·36 (0·64–2·89)	0·42
Mutation (unadjusted)	2·84 (1·29–6·69)	0·0090
Mutation (adjusted)*	2·79 (1·25–6·68)	0·012
>1 affected relative†	5·26 (1·83–15·30)	0·0022

0·1 0·3 1·0 3·0 8·0
Favours type L Favours type O

B Mutation status prediction (age <50 years)

	OR (95% CI)	p value
Non-white ethnicity	1·54 (0·56–4·52)	0·40
BMI >30 kg/m²	1·80 (0·75–4·58)	0·19
Current smoking	0·90 (0·32–2·47)	0·83
Age at menarche >12 years	1·13 (0·63–2·04)	0·69
OCP use >5 years	1·36 (0·74–2·52)	0·32
Current OCP use	1·26 (0·65–2·47)	0·50
Ever pregnant	1·22 (0·68–2·20)	0·51
HRT ever	0·91 (0·14–6·03)	0·92
Current combined hormone use	1·03 (0·50–2·13)	0·94
Combined hormone use >5 years	1·13 (0·62–2·03)	0·69
Community type O (unadjusted)	2·84 (1·29–6·69)	0·0090
Community type O (adjusted)‡	2·59 (1·16–6·18)	0·020

0·1 0·3 1·0 3·0 8·0
Favours control Favours case

C Community type prediction (age ≥50 years)

	OR (95% CI)	p value
Non-white ethnicity	1·64 (0·34–10·17)	0·55
BMI >30 kg/m²	1·27 (0·24–8·13)	0·78
Current smoking	0·67 (0·09–4·84)	0·68
Age at menarche >12 years	0·61 (0·16–2·21)	0·46
OCP use >5 years	0·56 (0·15–2·04)	0·37
Current OCP use	0·52 (0·08–3·07)	0·47
Ever pregnant	0·27 (0·00–3·65)	0·36
Postmenopausal state	3·80 (0·89–18·89)	0·070
HRT ever	1·65 (0·40–7·90)	0·49
Current combined hormone use	0·64 (0·09–4·61)	0·65
Combined hormone use >5 years	0·45 (0·12–1·64)	0·22
Mutation (unadjusted)	0·99 (0·28–3·61)	0·99

0·1 0·3 1·0 3·0 8·0
Favours type L Favours type O

D Mutation status prediction (age ≥50 years)

	OR (95% CI)	p value
Non-white ethnicity	1·34 (0·26–6·45)	0·72
BMI >30 kg/m²	1·80 (0·33–9·77)	0·48
Current smoking	1·88 (0·26–13·62)	0·51
Age at menarche >12 years	1·06 (0·30–3·89)	0·93
OCP use >5 years	1·63 (0·42–6·48)	0·48
Current OCP use	0·14 (0·00–1·44)	0·11
Ever pregnant	3·30 (0·25–467·03)	0·40
Postmenopausal state	1·19 (0·28–5·79)	0·82
HRT ever	1·17 (0·27–4·78)	0·82
Current combined hormone use	0·15 (0·00–1·61)	0·13
Combined hormone use >5 years	1·08 (0·28–4·02)	0·91
Community type O (unadjusted)	0·99 (0·28–3·61)	0·99

0·1 0·3 1·0 3·0 8·0
Favours control Favours case

Figure 2: **Effect of selected covariates on microbiota community type and risk of *BRCA1* mutation in the BRCA set**
(A) Prediction of microbiota community type in participants younger than 50 years. (B) Prediction of *BRCA1* mutation in participants younger than 50 years. (C) Prediction of microbiota community type in participants aged 50 years or older. (D) Prediction of *BRCA1* mutation in participants aged 50 years or older. Arrows represent truncated 95% CIs. More data are presented in the appendix (pp 13–17). OR=odds ratio. BMI=body-mass index. OCP=oral contraceptive pill. HRT=hormone replacement therapy. *Adjusted for pregnancy (ever). †Individuals with wild-type or mutated *BRCA1* who have more than one first-degree relative affected by any cancer. ‡Adjusted for postmenopausal state (appendix p 15).

was 39·5 days (SD 40·2; range 0–118). Time from collection to receipt did not differ between cases and controls (p=0·16), but time from receipt to pre-processing was shorter in cases with the *BRCA1* mutation than in controls (24·7 days [SD 30·5] *vs* 54·2 days [43·3]; p<0·0001). However, when time from receipt to pre-processing was taken individually as a predictor of community type L or type O in a logistic regression model in the BRCA set, the results were not significant (p=0·86 and p=0·49 for time from collection to receipt in those younger than 50 years and those aged 50 years or older, respectively; p=0·68 and p=0·71 for time from receipt to pre-processing in those younger than 50 years and those aged 50 years or older, respectively). Additionally, when time from collection to receipt and time from receipt to pre-processing were combined, the total time from sample collection to pre-processing differed significantly between cases with *BRCA1* mutations and controls (p=0·0002; data not shown). However, in the univariate logistic regression we did to confirm these results, the difference was no longer

significant (p=0·64 for participants younger than 50 years and p=0·592 for those aged 50 years or older).

The most important factor associated with a microbiota characterised by type L communities in the overall ovarian cancer set was age (p<0·0001), followed by menopausal status (p=0·0002; data not shown).

For the logistic regression model, values were missing only for epidemiological or clinical covariates; these participants were omitted (in each independent fitting, but not across the whole study) because they corresponded to a very small proportion for each variable (maximum 8·2%; appendix p 23).

With regard to the prevalence of community type L and type O microbiota, in the ovarian cancer set, women with benign gynaecological conditions and healthy volunteers did not differ according to a logistic regression model (p=0·49 for women younger than 50 years and p=0·97 for those aged 50 years or older). The corresponding p values were 0·99 and 0·32, respectively, in the BRCA set.

In women younger than 50 years in the ovarian cancer set, the only factor that was significantly predictive of
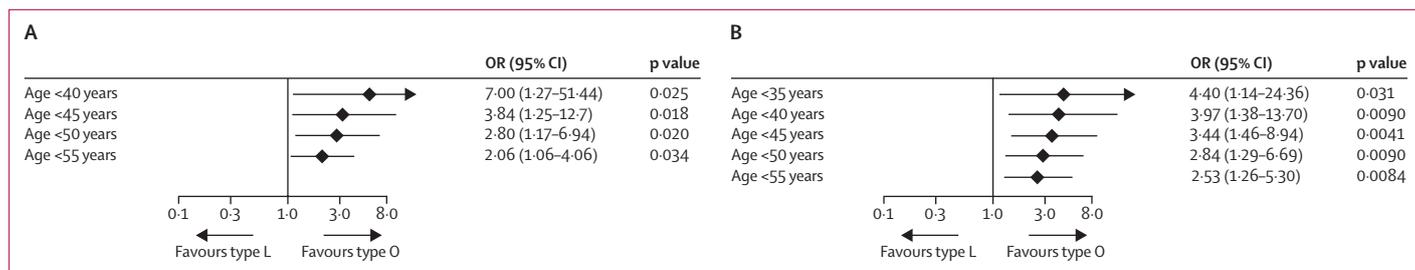
Figure 3: Prediction of microbiota community type for different age thresholds in the ovarian cancer set with ovarian cancer status as predictor (A), and in the BRCA set with mutation status as the predictor (B)

Arrows represent truncated 95% CIs. More data are presented in the appendix (pp 18–22). OR=odds ratio.

microbiota community status was the presence or absence of ovarian cancer: cases with ovarian cancer had a significantly increased prevalence of community type O microbiota than did age-matched controls (odds ratio [OR] 2·80 [95% CI 1·17–6·94], p=0·020; figure 1A; appendix pp 8–9). Conversely, in women younger than 50 years, several factors were significantly predictive of ovarian cancer status. In addition to current use of the oral contraceptive pill (OCP), duration of OCP use, ever use of hormone replacement therapy, current combined hormone use, and duration of combined hormone use, group community type was significantly predictive of ovarian cancer status, and remained significant after adjustment for additional covariates (OR 2·84 [95% CI 1·16–11·23]; p=0·023; figure 1B; appendix p 10). Both purely additive models and those with interaction terms were tested. Here, we report the additive models because only these showed significant results (data not shown for models with interaction terms).

In women aged 50 years or older in the ovarian cancer set, there was a tendency towards a higher prevalence of type O communities than type L communities in the cervicovaginal samples (table 1). Women who had been on the OCP for more than 5 years or on combined hormones for more than 5 years were more likely to have community type L microbiota than type O microbiota, and had a lower risk of developing ovarian cancer than those who had been on the OCP or using combined hormones for less than 5 years (figure 1C, 1D; appendix pp 11–12). The only additional factor that was significantly predictive of community status in women aged 50 years or older was postmenopausal status (which was predictive of community type O microbiota status; figure 1C) and the only additional factors predictive of ovarian cancer status in these patients were current smoking, ever pregnant, and ever use of hormone replacement therapy (figure 1D).

We assessed whether a type O cervicovaginal microbiota community was associated with an ovarian cancer predisposition in the 109 age-matched BRCA1 mutation carriers and 111 age-matched controls who had available microbiome data. In this BRCA set, there was a higher preponderance of a type O cervicovaginal microbiota community in women aged 50 years or older, whereas most of the women aged younger than 50 years had a type L microbiota community (table 2). In women younger than 50 years, both pregnancy (ever) and the presence of the BRCA1 mutation were associated with an increased prevalence of community type O microbiota (figure 2A). After adjustment for pregnancy, the OR for having cervicovaginal microbiota dominated by community type O was 2·79 (95% CI 1·25–6·68; p=0·012) in BRCA1 mutation carriers compared with those who did not have this mutation (figure 2A; appendix pp 13–14).

In the BRCA set, 34 BRCA1 mutation carriers had more than one first-degree family member who had been affected by any cancer—a condition that further increases the future likelihood of developing ovarian cancer by 1·8 times.[21] The OR for having a community type O microbiota was 5·26 (95% CI 1·83–15·30; p=0·0022) in these 34 women compared with those without a BRCA1 mutation (figure 2A). The BRCA set included 45 women younger than 30 years of age, of whom 22 were mutation carriers and 23 were controls. Six (27%) of the 22 with a BRCA1 mutation had microbiota dominated by community type O, compared with none of the women without the mutation.

Among women younger than 50 years in the BRCA set, microbiota community type was the only significant predictor of BRCA1 mutation status: the OR for being a BRCA1 mutation carrier was 2·84 (95% CI 1·29–6·69; p=0·0090) in women with community type O microbiota compared with those with type L microbiota (figure 2B; appendix p 15). Among the 40 women who were aged 50 years or older in the BRCA set, we noted no significant associations between clinical variables and community type or BRCA1 mutation status (figure 2C, 2D; appendix pp 16–17).

Finally, we assessed whether the association between community type O microbiota and ovarian cancer or BRCA1 mutation status depended on our chosen age cutoff. To this end, the analyses for each of the covariates collected for this study were repeated in both the ovarian cancer and BRCA sets with cutoffs of 40 years, 45 years, and 55 years (appendix pp 18–22), and additionally with a cutoff of 35 years for the BRCA set, which had sufficient numbers of cases and controls in this age group (appendix p 20). In both sets, the lower the age threshold,

the stronger the association in the younger age group between community type O microbiota and ovarian cancer status (figure 3A) or *BRCA1* mutation status (figure 3B).

Discussion

In this study, we showed that the prevalence of women with non-lactobacilli-dominated cervicovaginal micro-biomes was higher both in patients with ovarian cancer and in women with *BRCA1* mutations who had yet to develop cancer, compared with age-matched healthy women and women without *BRCA1* or *BRCA2* mutations, respectively.

To assess the cervicovaginal microbiome in these participants, we chose to specifically collect cervicovaginal samples (as opposed to vaginal samples, which are rather similar[22]) because we wanted to assess samples from an easy-to-access anatomical area that serves as a barrier for ascending infections affecting the fallopian tube (ie, the organ from which ovarian cancer originates). The rationale for the division of the samples into two age groups—those younger than 50 years and those aged 50 years or older—was premised on two considerations. In previous work, only volunteers younger than the so-called reproductive age—ie, 45 years—were recruited.[7] Had we used age 45 years as the cutoff for the ovarian cancer set, the number of ovarian cancer cases included would have been lower, because the disease is more prevalent in older women than in those younger than 45 years. Conversely, in several studies,[8,9,23] the loss of *Lactobacillus* species from the cervicovaginal microbiota—and, consequently, an increase of community type O microbiota—was associated with decreases in oestrogen concentrations in postmenopausal women (in association with glycogen decreases). Indeed, in our ovarian cancer set, the higher prevalence of older than younger women with community type O microbiota is consistent with previous reports.[7] In view of the challenges associated with the design of a protocol for diseases that are more prevalent in one specific age group than in others (as shown by Walther-Antonio and co-workers[4]), we chose an age-matched design for our study. Nevertheless, a division of the cohort into age groups was still necessary to avoid the strong confounding effect of age as a major contributor to ovarian cancer and the microbiome.

Our data suggest that the presence of ovarian cancer or factors known to increase the risk of ovarian cancer (ie, older age or *BRCA1* germline mutation) are associated with the absence of a protective community type L cervicovaginal microbiota. The finding that the prevalence of community type L microbiota was lower in patients with ovarian cancer than in healthy controls, irrespective of disease stage, suggests that this reduced prevalence of type L microbiota could be causal rather than consequential. Although our data provide no direct proof that community type O microbial species dominance (and thereby a potential reduced protection

from ascending infections) has a causal role in the development of ovarian cancer, several previously reported findings potentially support this hypothesis. First, an analysis[24] of the microbiota of ovarian cancer tissue samples and matched non-cancerous tissue showed that two bacterial phyla, namely *Proteobacteria* and *Firmicutes*, were predominant in the cancer samples but not in the control samples. Second, inflammation caused by such agents as *Proteobacteria* and *Firmicutes*, and the subsequent release of bacterial toxins such as colibactin and cytolethal distending toxin, could directly damage the cellular DNA by causing double-stranded breaks and thereby activating the DNA damage checkpoint pathway.[25] Third, the protective mechanism of interventions that reduce ovarian cancer risk could be mediated, at least partly, by modulation of the microbiota. Tubal ligation[26] could potentially limit access of ascending bacteria to the fallopian tube and ovaries. Long-term use of a combined OCP could reduce the risk of ovarian cancer[27] by supporting the colonisation of the vagina with beneficial lactic-acid-producing lactobacilli.[28] In our study, the finding that in women aged 50 years or older, use of the OCP for more than 5 years is associated with community type L microbiota and reduced risk of ovarian cancer raises the question of whether long-term use of the OCP could be associated with a delay in age-triggered predominance of community type O in the cervicovaginal microbiota. Whether the association between long-term OCP use and beneficial cervicovaginal microbiota in women aged 50 years or older is merely a surrogate marker for reduced disease risk or whether OCP use has a causal effect on the preservation of a beneficial cervicovaginal microbiota, which in turn protects against ovarian cancer, remains to be established.

Several lines of additional evidence support the view that, in *BRCA1* mutation carriers, cell-non-autonomous (systemic) factors trigger the growth of community type O species in the vagina. We have previously shown that *BRCA1* or *BRCA2* mutation carriers have increased concentrations of progesterone during the luteal phase of their menstrual cycles.[29] Vaginal lactobacilli metabolise glucose and maltose—the breakdown products of human glycogen (which is substantially reduced in postmeno-pausal women)—through fermentation, thereby produc-ing lactic acid.[8] Progesterone is known to reduce vaginal glycogen concentrations, resulting in an environment that is less conducive to the growth of community type L species. Thus, the reduced prevalence of community type L cervicovaginal microbiota in both women with ovarian cancer and *BRCA1* mutation carriers younger than 50 years without a diagnosis of ovarian cancer could possibly be a reflection of an accumulating progesterone–oestrogen imbalance.

Our study has several potential limitations. Most of the women included were white. The microbiome (both in general and specifically the vaginal microbiome)[7] varies substantially by ethnicity[30] and region,[31] and therefore

our analyses should be repeated in additional cohorts. Another possible limitation is that we did not test people in the ovarian cancer set for *BRCA* germline mutations. However, given that only 10–12% of women diagnosed with invasive ovarian cancer have germline mutations in *BRCA1* or *BRCA2*,[32] it is unlikely that the presence of mutations would have inflated the ORs substantially. As is the case in most studies involving biosamples, the time between sample collection and processing varied among patients. However, all samples were directly transferred to PreservCyt media (which is designed to preserve cellular and subcellular components and prevent microbial growth), and neither the transit time nor the time to pre-processing had any effect on composition of the microbiota community. Thus, it is unlikely that a storage-triggered bias could have substantially affected our results.

In summary, we found that community type O bacterial compositions dominate the cervicovaginal microbiota in women aged 50 years or older. Furthermore, the presence of ovarian cancer or a germline *BRCA1* mutation favoured a composition of the cervicovaginal microbiota (ie, a type O community) that is more commonly seen in older than in younger women. Although our findings provide no direct evidence that this seemingly premature ageing of the cervicovaginal microbiome is causally involved in the genesis of ovarian cancer, evidence showing an altered microbiome in ovarian cancer samples compared with control fallopian tube samples[33] appears to support this line of reasoning. Our findings warrant further detailed analyses of the vaginal microbiome, especially in high-risk women (ie, with *BRCA1* mutation) younger than 50 years, and in women with community type O microbiota (particularly young white women). However, before a clinical trial is initiated, our findings need to be further validated, and longitudinal assessments of the microbiome in mutation carriers will be essential. The application of vaginal suppositories containing a combination of live lactobacilli and oestriol normalises the vaginal microbiota.[34] Establishing whether a continuous community type L cervicovaginal microbiome could reduce the risk of ovarian cancer should be a priority for further studies in this area.

Contributors
MW conceived and designed the study, and drafted the Article. NRN did the statistical analyses, produced the figures, and interpreted the data in collaboration with MW. AL and TP did the DNA extraction and sequenced the microbiomes. All other authors contributed to data acquisition and interpretation, and critically reviewed and approved the Article.

Declaration of interests
AL and TP are employees of Eurofins, which offers 16S rRNA gene sequencing as a service. NC reports personal fees from Roche, Pharmamar, AstraZeneca, Clovis, Tesaro, Pfizer, Takeda, and Biocad. All other authors declare no competing interests.

Data sharing
The original data including all clinical, epidemiological, and microbiota data used in this work will be made available upon request. Requests should be directed to the corresponding author.

References
1 Garrett WS. Cancer and the microbiota. *Science* 2015; **348:** 80–86.
2 McQuade JL, Daniel CR, Helmink BA, Wargo JA. Modulating the microbiome to improve therapeutic response in cancer. *Lancet Oncol* 2019; **20:** e77–91.
3 Bucevic V, Situm M, Chow CT, Chan LS, Roje B, Terzic J. The urinary microbiome associated with bladder cancer. *Sci Rep* 2018; **8:** 12157.
4 Walther-Antonio MR, Chen J, Multinu F, et al. Potential contribution of the uterine microbiome in the development of endometrial cancer. *Genome Med* 2016; **8:** 122.
5 Jayson GC, Kohn EC, Kitchener HC, Ledermann JA. Ovarian cancer. *Lancet* 2014; **384:** 1376–88.
6 Lin H-W, Tu Y-Y, Lin SY, et al. Risk of ovarian cancer in women with pelvic inflammatory disease: a population-based study. *Lancet Oncol* 2011; **12:** 900–04.
7 Ravel J, Gajer P, Abdo Z, et al. Vaginal microbiome of reproductive-age women. *Proc Natl Acad Sci USA* 2011; **108** (suppl 1)**:** 4680–87.
8 Smith SB, Ravel J. The vaginal microbiota, host defence and reproductive physiology. *J Physiol* 2017; **595:** 451–63.
9 van de Wijgert JH, Borgdorff H, Verhelst R, et al. The vaginal microbiota: what have we learned after a decade of molecular characterization? *PLoS One* 2014; **9:** e105998.
10 Reid G. Cervicovaginal microbiomes—threats and possibilities. *Trends Endocrinol Metab* 2016; **27:** 446–54.
11 Mavaddat N, Peock S, Frost D, et al. Cancer risks for *BRCA1* and *BRCA2* mutation carriers: results from prospective analysis of EMBRACE. *J Natl Cancer Inst* 2013; **105:** 812–22.
12 Mahe F, Mayor J, Bunge J, et al. Comparing high-throughput platforms for sequencing the V4 region of SSU-rDNA in environmental microbial eukaryotic diversity surveys. *J Eukaryot Microbiol* 2015; **62:** 338–45.
13 Klindworth A, Pruesse E, Schweer T, et al. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res* 2013; **41:** e1.
14 Magoc T, Salzberg SL. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 2011; **27:** 2957–63.
15 Li W, Godzik A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 2006; **22:** 1658–59.
16 Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 2011; **27:** 2194–200.
17 Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol* 1990; **215:** 403–10.
18 Cole JR, Wang Q, Cardenas E, et al. The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Res* 2009; **37:** D141–45.
19 Federhen S. The NCBI taxonomy database. *Nucleic Acids Res* 2012; **40:** D136–43.
20 Pawitan Y, Michiels S, Koscielny S, Gusnanto A, Ploner A. False discovery rate, sensitivity and sample size for microarray studies. *Bioinformatics* 2005; **21:** 3017–24.
21 Kuchenbaecker KB, Hopper JL, Barnes DR, et al. Risks of breast, ovarian, and contralateral breast cancer for *BRCA1* and *BRCA2* mutation carriers. *JAMA* 2017; **317:** 2402–16.

22   Chen C, Song X, Wei W, et al. The microbiota continuum along the female reproductive tract and its relation to uterine-related diseases. *Nat Commun* 2017; **8:** 875.

23   Brotman RM, Shardell MD, Gajer P, et al. Interplay between the temporal dynamics of the vaginal microbiota and human papillomavirus detection. *J Infect Dis* 2014; **210:** 1723–33.

24   Banerjee S, Tian T, Wei Z, et al. The ovarian cancer oncobiome. *Oncotarget* 2017; **8:** 36225–45.

25   Mert I, Walther-Antonio M, Mariani A. Case for a role of the microbiome in gynecologic cancers: clinician's perspective. *J Obstet Gynaecol Res* 2018; **44:** 1693–704.

26   Cibula D, Widschwendter M, Majek O, Dusek L. Tubal ligation and the risk of ovarian cancer: review and meta-analysis. *Hum Reprod Update* 2011; **17:** 55–67.

27   Beral V, Doll R, Hermon C, Peto R, Reeves G. Ovarian cancer and oral contraceptives: collaborative reanalysis of data from 45 epidemiological studies including 23 257 women with ovarian cancer and 87 303 controls. *Lancet* 2008; **371:** 303–14.

28   Brooks JP, Edwards DJ, Blithe DL, et al. Effects of combined oral contraceptives, depot medroxyprogesterone acetate and the levonorgestrel-releasing intrauterine system on the vaginal microbiome. *Contraception* 2017; **95:** 405–13.

29   Widschwendter M, Rosenthal AN, Philpott S, et al. The sex hormone system in carriers of BRCA1/2 mutations: a case-control study. *Lancet Oncol* 2013; **14:** 1226–32.

30   Gaulke CA, Sharpton TJ. The influence of ethnicity and geography on human gut microbiome composition. *Nat Med* 2018; **24:** 1495–96.

31   He Y, Wu W, Zheng HM, et al. Regional variation limits applications of healthy gut microbiome reference ranges and disease models. *Nat Med* 2018; **24:** 1532–35.

32   Risch HA, McLaughlin JR, Cole DE, et al. Prevalence and penetrance of germline *BRCA1* and *BRCA2* mutations in a population series of 649 women with ovarian cancer. *Am J Hum Genet* 2001; **68:** 700–10.

33   Zhou B, Sun C, Huang J, et al. The biodiversity composition of microbiome in ovarian carcinoma patients. *Sci Rep* 2019; **9:** 1691.

34   Ozkinay E, Terek MC, Yayci M, Kaiser R, Grob P, Tuncay G. The effectiveness of live lactobacilli in combination with low dose oestriol (Gynoflor) to restore the vaginal flora after treatment of vaginal infections. *BJOG* 2005; **112:** 234–40.