MICROBES OF THE FEMALE GENITAL TRACT, INFLAMMATION AND OVARIAN CANCER

Dr Eilbhe Cecile Whelan

Imperial College London

Metabolism, Digestion and Reproduction, Imperial College London

A thesis submitted for the Degree of MD (Research)

Supervisors:

Professor Maria Kyrgiou

Professor David MacIntyre

Professor Philip Bennett

Professor Julian Marchesi

Declaration of Originality

I declare that all work presented in this thesis is entirely my own unless stated otherwise. Published and unpublished data and articles, referred to throughout this thesis, are duly cited within the body of the text and referenced in the bibliography in accordance with College guidance.

London,
March 2023

Eilbhe Whelan

Copyright Declaration

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial 4.0 International License (CC BY-NC).

Under this license, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose.

When reusing or sharing this work, ensure you make the license terms clear to others by naming the license and linking to the license text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes.

Please seek permission from the copyright holder for uses of this work that are not included in the license or permitted under UK Copyright Law

ABSTRACT

Background:

Eighty-five percent of cases of ovarian cancer are sporadic and the mechanisms behind ovarian oncogenesis are unclear. Several putative modifiable risk factors have been suggested but robust evidence is lacking. Exploratory studies in ovarian cancer have demonstrated cervicovaginal dysbiosis both in women with ovarian cancer and those at risk of ovarian cancer e.g., carrying *BRCA* mutation.

Aims:

The first aim was to critically appraise the strength of epidemiological evidence for associations between non-genetic risk factors and ovarian cancer. The second aim was to investigate if genuine bacterial signals from low bacterial biomass sites of the upper genital tract can be detected above background contaminants and if so, establish whether a bacterial continuum is present along throughout the genital tract of ovarian cancer patients and patients with benign pathology. The third aim was to test the hypothesis that ovarian cancer patients harbour a distinctive bacterial fingerprint in their genital tract compared to benign controls.

Results:

An umbrella review of the literature identified six associations graded with strong evidence which increased the risk of developing ovarian cancer, and one which reduced the risk.

Metataxonomic profiling of bacterial communities of the reproductive tract of ovarian cancer (n=19) and control patients (n=12) along with controls indicated that approximately 20% of recovered sequencing data could be attributed to contamination. Cervicovaginal dysbiosis was more prevalent in the ovarian cancer cohort. The benign cohort all demonstrated a continuum of bacterial species with a similar trend noted in 80% of malignant patients. Differentially abundant taxa were identified at all reproductive tract sites when comparing the benign and malignant cohorts.

Discussion:

The results of this descriptive exploratory study contributes to the hypothesis that the bacterial signature of the female genital tract may play a role in ovarian oncogenesis and warrants extensive, systematic investigation to further explore host-microbiome interactions.

Table of Contents

Declaration of Originality 2
Copyright Declaration 2
Abstract 3
Table of Contents 4
List of figures 8
List of tables 10
List of Abbreviations 12
Acknowledgements 14

CHAPTER ONE: INTRODUCTION 17
1.1 EPIDEMIOLOGY OF OVARIAN CANCER 18
1.2 RISK FACTORS FOR OVARIAN CANCER 18
1.2.1 Genetic predisposition 19
1.2.2 Non-genetic risk factors 19
1.3 OVARIAN CANCER PREVENTION 26
1.3.1 Risk reducing bilateral salpingo-oophrectomy (RRBSO): primary prevention of ovarian cancer 26
1.3.2 Screening for ovarian cancer: secondary prevention of ovarian cancer 26
1.4 HISTOPATHOLOGY OF OVARIAN CANCER 29
1.4.1 Epithelial tumours 29
1.4.2 Fallopian tube and primary peritoneal cancers 32
1.5 STAGING AND GRADING OF OVARIAN CANCER 33
1.6 PATHOGENESIS OF OVARIAN CANCER 34
1.6.1 Molecular pathogenesis of ovarian cancer 34
1.6.2 Proinflammatory conditions and ovarian cancer 35
1.7 TREATMENT OF OVARIAN CANCER 35
1.7.1 Cytoreductive surgery 36
1.7.2 Adjuvant chemotherapy 36
1.8 THE HUMAN MICROBIOME 36
1.8.1 The microbiome of the female genital tract 37
1.9 HYPOTHESIS AND AIMS 39
1.9.1 Rationale 39
1.9.2 Hypothesis 41

1.9.3 Aims 41

CHAPTER TWO: MATERIALS AND METHODS 43
2.1 MATERIALS 44
2.2 METHODS FOR THE UMBRELLA REVIEW 45
2.2.1 Literature search 45
2.2.2 Eligibility criteria 46
2.2.3 Evaluation of the strength of evidence 47
2.2.4 Evaluation of the quality of included meta-analyses 47
2.2.5 Data Extraction 47
2.2.6 Data Analysis 48
2.3 METHODS FOR THE ANALYSIS OF THE FEMALE GENITAL TRACT MICROBIOTA 50
2.3.1 Study population 50
2.3.2 Sample collection and processing 51
2.3.3 Microbiome bacterial DNA extraction 51
2.3.4 16S rRNA polymerase chain reaction (PCR) 52
2.3.5 Sequencing of bacterial 16S rRNA genes 53
2.3.6 Processing of sequence data 53
2.3.7 Statistical analyses 54
2.3.8 Contamination considerations in microbiome research 56

CHAPTER THREE: RISK FACTORS FOR OVARIAN CANCER: AN UMBRELLA REVIEW OF THE LITERATURE 57
3.1 ABSTRACT 58
3.1.1 Background 58
3.1.2 Materials and methods 58
3.1.3 Results 58
3.1.4 Conclusions 58
3.2 AIMS 59
3.3 INTRODUCTION 59
3.4 RESULTS 60
3.4.1 Characteristics of meta-analyses 60
3.4.2 Summary effect size 63
3.4.3 Heterogeneity between studies 63
3.4.4 Small study effects 63
3.4.5 Excess significance 64

3.4.6 Quality assessment 64
3.4.7 Grading the evidence 74
3.4.8 Sensitivity Analysis 87

3.5 DISCUSSION 95
3.6 STRENGTHS AND LIMITATIONS 98
3.7 CONCLUSIONS 98
3.8 STATEMENT OF CONTRIBUTION 99

CHAPTER FOUR: INVESTIGATING THE PRESENCE OF A GENUINE BACTERIAL SIGNATURE IN THE FEMALE UPPER GENITAL TRACT AND THE PRESENCE OF A BACTERIAL CONTINUUM 100
4.1 INTRODUCTION 101
4.2 AIMS 102
4.3 RESULTS 102

4.3.1 Study populations characteristics 102
4.3.2 Contamination considerations in microbiome research 103
4.3.3 The upper genital tract contains bacterial metataxonomic signatures above that of background controls 113
4.3.4 Patients with malignant and benign conditions displayed common microbes throughout their genital tract. 117
4.3.5 The endometrium and UGT shared a greater proportion of bacterial species than the LGT and the endometrium. 119

4.4 DISCUSSION 121
4.5 STRENGTHS AND LIMITATIONS 124
4.6 CONCLUSIONS 124
4.7 STATEMENT OF CONTRIBUTION 125

CHAPTER FIVE: COMPARISON OF FEMALE GENITAL TRACT BACTERIAL PROFILES IN PATIENTS WITH OVARIAN CANCER AND BENIGN CONTROLS 126
5.1 INTRODUCTION 127
5.2 AIMS 128
5.3 RESULTS 128

5.3.1 Lactobacillus species depletion of the female genital tract is associated with ovarian cancer. 128
5.3.2 Each anatomical site was noted to display distinguishing species between the benign and malignant samples 138
5.3.3 Ovarian cancer samples displayed higher α-diversity than the benign controls at both species and genus level in the lower genital tract. 141

5.3 DISCUSSION 142

5.4 STRENGTHS AND LIMITATIONS 145
5.5 CONCLUSIONS 145
5.6 STATEMENT OF CONTRIBUTION 146

CHAPTER SIX: CONCLUSIONS 147
6.1 SUMMARY OF HYPOTHESIS 148
6.2 SUMMARY OF FINDINGS 149
6.2.1 Umbrella review 149
6.2.2 Female genital tract presence 149
6.2.3 Female genital tract composition 150
6.3 PROPOSED MECHANISM 151
6.4 TRANSLATIONAL VALUE 152
6.5 FUTURE DIRECTION 153

REFERENCES 155
APPENDIX 169
PUBLICATIONS 193
PRESENTATIONS 193

LIST OF FIGURES

FIGURE 1.1 Summary of causes of a raised CA125

FIGURE 1.2 Common phyla in the female genital tract

FIGURE 2.1 Description of selection of statistical analyses used in this thesis

FIGURE 2.2 Flow diagram of selection of meta-analyses on risk factors for ovarian cancer

FIGURE 3.1 Flow diagram of selection of meta-analyses on risk factors for ovarian cancer

FIGURE 4.1 Flow diagram of sample selection

FIGURE 4.2 Venn diagram demonstrating the results of the three different thresholds for identifying suspected contaminated DNA sequences.

FIGURE 4.3 Difference in read counts pre and post application of of decontam package per site in the benign and cancer groups, and all samples

FIGURE 4.4 Depiction of the proportion of contaminants in each collection of samples by anatomical site and control group (a) and individual samples per site (b)

FIGURE 4.5 Correlation analysis of upper and lower endometrial samples at genera and species level

FIGURE 4.6 Univariate comparison analysis of laparoscopic and transabdominal samples by body site

FIGURE 4.7 Percentage of patients from the cancer cohort who displayed a common bacterium throughout the genital tract

FIGURE 4.8 Percentage of patients from the benign cohort who displayed a common bacterium throughout the genital tract

FIGURE 4.9 Venn diagrams illustrating the number of shared species throughout the genital tract of a benign and malignant patient

FIGURE 4.10 Venn diagram of the number of bacterial species identified which are shared between the lower genital tract, endometrium and upper genital tract

FIGURE 4.11 Number of species that overlap at different anatomical sites and compartments. All samples per site included for analysis

FIGURE 5.1 Hierarchal clustering analysis of all vaginal samples using both Ward linkage and Euclidean distance and the VALENCIA method at genus level

FIGURE 5.2 Hierarchal clustering analysis of all cervical samples using both Ward linkage and Euclidean distance and the VALENCIA method at genus level

FIGURE 5.3 Hierarchal clustering analysis of all endocervical samples using Ward linkage and Euclidean distance and the VALENCIA method at genus level

FIGURE 5.4 Hierarchal clustering analysis of all fundal endometrial samples using Ward linkage and Euclidean distance

FIGURE 5.5 Hierarchal clustering of fallopian tube samples using Ward linkage and Euclidean distance

FIGURE 5.6 Hierarchal clustering of ovarian samples using Ward linkage and Euclidean distance

FIGURE 5.7 Univariate analysis of benign and malignant vaginal samples.
FIGURE 5.8 Univariate analysis of benign and malignant cervical samples
FIGURE 5.9 Univariate analysis of benign and malignant endometrial samples
FIGURE 5.10 Univariate analysis of benign and malignant fallopian tube samples
FIGURE 5.11 Proportion of bacteria in each cohort at phylum level
FIGURE 5.12 Univariate analysis of benign and malignant ovarian samples
FIGURE 5.13 α and β diversity of vaginal samples at species and genus level

LIST OF TABLES

TABLE 1.1 Details of studies investigating the presence of microbes and ovarian cancer, and PID and ovarian cancer

TABLE 1.2 Ultrasonographic features of benign and malignant ovarian masses as per the IOTA group

TABLE 1.3 Classification of ovarian tumours by Kurman and Shih[1]

TABLE 1.4 2014 FIGO criteria for staging cancers of the ovaries, fallopian tubes and peritoneum

TABLE 2.1 List of materials and reagents used

TABLE 2.2 Search algorithm for ovarian cancer umbrella review

TABLE 2.3 Details of criteria for evidence grading

TABLE 2.4 Bacterial PCR Mastermix (Total 50µl per reaction)

TABLE 2.5 PCR conditions for 16S rRNA primers

TABLE 3.1 Details of risk factors assessed

TABLE 3.2 Description of 40 meta- analyses results investigating risk factors associated with ovarian cancer incidence or mortality –cohort studies.

TABLE 3.3 Evaluation of heterogeneity, small study effects and excess significance bias in the 40 meta-analyses investigating the risk factors associated with ovarian cancer incidence or mortality[§] –cohort studies.

TABLE 3.4 A Measurement Tool to Assess Systematic Reviews 2 (AMSTAR 2) Summary quality assessment for all included systematic reviews

TABLE 3.5 Summary of evidence grading for meta-analysis of risk factors associated with ovarian cancer incidence or mortality– cohort studies only*

TABLE 3.6 Details of evidence grading for meta-analysis of risk factors for ovarian cancer incidence or mortality – only cohort studies included*

TABLE 3.7 Details of evidence grading for meta-analyses of risk factors for ovarian cancer incidence or mortality[κ] – all statistically significant study types included*

TABLE 3.8 Sensitivity analysis using credibility ceilings when the association is non-significant of the 40 studies investigating the risk factors associated with incidence or mortality[§] of ovarian cancer - only cohort studies included.

TABLE 3.9 Excluded duplicate studies and studies selected in their place to be included in the analysis

TABLE 4.1 Patient characteristics at baseline by histology status

TABLE 4.2 Repeating patterns of sequence reads across all samples indicates carry over of species DNA from the first samples and subsequent contamination of samples further along the processing line.

TABLE 4.3 Species enriched in upper genital tract sites versus controls

TABLE 4.4 16S rRNA sequence reads by anatomical body site and histology

TABLE 4.5 The relative abundance of the shared species found throughout the reproductive tract as a percentage of the total number of unique species per patient.

TABLE 4.6 Percentage relative abundance of common species between anatomical sites when compared directly.

TABLE 5.1 Results of hierarchal clustering of vaginal samples using a) Euclidean distance and Ward linkage and then using b) VALENCIA method

TABLE 5.2 Results of hierarchal clustering of cervical samples using a. Euclidean distance and Ward linkage and then b. VALENCIA method

TABLE 5.3 Results of hierarchal clustering of endocervical samples using Euclidean distance and Ward linkage.

TABLE 5.4 Results of hierarchal clustering of fundal endometrial samples using Euclidean distance and Ward linkage

TABLE 5.5 Results of hierarchal clustering analysis of fallopian tube samples using Euclidean distance and Ward linkage.

TABLE 5.6 Results of hierarchal clustering of ovarian samples using Euclidean distance and Ward linkage

List of Abbreviations

BMI	Body mass index
BOT	Borderline ovarian tumours
BS	Bilateral salpingectomy
BSO	Bilateral salpingo-oophrectomy
BRCA	BReast CAncer gene
BV	Bacterial vaginosis
CCC	Clear cell carcinoma
CI	Confidence interval
CIN	Cervical intraepithelial neoplasia
CST	Community state type
DNA	Deoxyribonucleic acid
EAOC	Endometriosis- associated ovarian cancer
EC	Endometrioid carcinoma
ECM	Extra cellular matrix
EDTA	Ethylenediaminetetraacetic Acid
FIGO	International Federation of Gynecology and Obstetrics
FDR	False discovery rate
GWAS	Genome wide association studies
HCA	Hierarchal clustering analysis
HGSC	High grade serous cancer
HIPEC	Hyperthermic intraperitoneal chemotherapy
HMP	Human Microbiome Project
HRT	Hormone replacement therapy
HPV	Human papilloma virus
IARC	International Agency for Research on Cancer
IOTA	International Ovarian Tumour Analysis group
IVF	Invitro fertilisation
KRAS	Kirsten rat sarcoma
LGE	Low grade endometrioid
LGSC	Low grade serous carcinoma
LGT	Lower genital tract
LPS	Lipopolysaccharide
MC	Mucinous carcinoma
NF - κB	Nuclear factor κB

NGS	Next generation sequencing
OCP	Oral contraceptive pill
OTU	Operational taxonomic unit
PAMP	Pathogen-associated molecular patterns
PARP	Poly (ADP- ribose) polymerase
PBS	Phosphate buffered saline
PCA	Principal component analysis
PCR	Polymerase chain reaction
PDS	Primary debulking surgery
PID	Pelvic inflammatory disease
PRR	Pattern recognition receptor
PTEN	Phosphatase and tensin homolog
RCOG	Royal College of Obstetrics and Gynaecology
RMI	Risk of malignancy index
RNA	Ribonucleic acid
RRBSO	Risk reducing bilateral salpingo-oophrectomy
SBT	Serous borderline tumour
SD	Standard deviation
SET	Solid, endometrioid-like, and transitional
SHBG	Sex- hormone binding globulin
SLPI	Secretory leukocyte protease inhibitor
Spp.	Species (plural)
STIC	Serous tubal intraepithelial carcinoma
TLR	Toll like receptor
TNF	Tumour necrosis factor
TVS	Transvaginal scan
UGT	Upper genital tract
UKTOCS	UK Collaborative Trial of Ovarian Cancer Screening
UTI	Urinary tract infection
UV	Ultraviolet
WHO	World Health Organisation

ACKNOWLEDGEMENTS

Without doubt, the completion of this research project and writing of this thesis has been the most challenging endeavour of my life to date. There have been many times when I have faltered but was able to persevere due to the help and encouragement of my family, friends and colleagues. Research is above all a collaborative effort, and I am fortunate to have both personal and professional support networks to lean on.

First, I would like to thank the patients who agreed to participate in this study. Without their willingness to help others who might find themselves in a similar situation, there would be no research. Thank you for agreeing to speak to an over-enthusiastic clinical fellow during what was an undoubtably tricky journey.

Thank you to Professor Mara Kyrgiou who agreed to take on an unknown transplant from Ireland via Norfolk as a research fellow. Your capacity to keep innumerable plates in the air whilst also conceiving and executing novel and vital projects is astonishing. Thank you for your support, guidance and help. You are an inspirational mentor and brilliant clinician- academic.

Thank you to Professor David MacIntyre for your infinite patience, encouragement and many recorded virtual meetings! Your guidance and advice, from the minutiae of working an excel spreadsheet to demonstrating how to conceptualise a research topic, will serve me throughout the rest of my career. You are a truly inspiring scientist, and I feel incredibly fortunate to have learnt from you.

To Professors Philip Bennett and Julian Marchesi, thank you for being generous with your time and contributions to my research at several points over the years, and for your timely and focused reviews of manuscripts.

Thank you to Mr Nikos Burbos who set this MD Research train in motion by introducing me to Mara, and to the rest of the Gynaecological Oncology team at the Norfolk and Norwich University Hospital who have all continued to encourage and support me over the years.

Thank you to;

Dr Yun Lee, who kindly leant her time to teach me how to perform the microbiome DNA extraction technique, and how to be an exacting scientist! All in the lab group on the 3rd floor who were unfailingly friendly and helpful to a very grateful but very green clinician.

Dr Ilkka Kalliala, who was eternally patient in explaining the ins and outs of an umbrella review and always quick to respond to my many queries.

To everyone in the Gynaecology Oncology outpatient clinics, operating theatres and pathology department who facilitated the recruitment of patients and obtaining of samples.

Mr Richard Smith and Mr James Kinross for your assistance and input to my Early and Late Stage Reviews.

Infinite thank yous to fellow clinician researcher, and now lifelong friend, Dr Olivia Raglan whose support, encouragement and generous hosting kept me going- conferences, weddings, babies; we did it all!

To Dr Anita Mitra, who showed me the ropes at the beginning of my study and Dr Anita Semertzidou who both acted as a sounding board and fielded many questions and queries, I greatly appreciate your input.

I am privileged beyond words to be surrounded by several groups of incredible friends who have been alternately shoulders to cry on, confidantes and cheerleaders. I am lucky that there are too many to mention, you know who you are.

Our dear friends, Lou and Keith, deserve special mention for opening their home to us initially on our own, and then in the 'commune', after Hazel was born amid my research. Your unspoken yet unwavering support is never taken for granted. You will never know how much both Jamie and I value our friendship.

My parents never for one second doubted that I could achieve anything I set my mind to. Thank you for providing me with every opportunity I could have ever wished for and for your endless support, encouragement and love. I promise I will finish my training soon. Ish.
Thank you to my siblings who never fail to bring me down to earth whilst quietly pushing me forwards.

Hazel and Meabh bookended my MD programme. Hazel arrived with a bit more fanfare and a bit less notice that we would have anticipated. You successfully redrew the plans of my life and were the one to turn me into a mother. Your giggle and singing makes the world spin. Meabh, you are part of the final chapter of this MD. Your attitude, joy and love of kicking a ball never fails to bring a smile to my face.

I am forever grateful and proud to be your mother. Thank you for your patience whilst I finish this thesis.

Jamie, my husband, has been unwavering in his support and love over the past six years of this research. We have not taken the easiest road, but our walls have held fast. I marvel at your capacity for compromise and sacrifice and simply could not have completed this MD without your help and guidance in all aspects. I am so proud of our family and I cannot wait for our first holiday without my laptop. Thank you.

'Never fear. We may let the scaffolds fall
Confident that we have built our wall'

Seamus Heaney 1939-2013

CHAPTER ONE: INTRODUCTION

1.1 EPIDEMIOLOGY OF OVARIAN CANCER

Ovarian cancer is the sixth most common cancer in women in the UK[2]. It is known as the 'silent killer' due to its pernicious and insidious nature. It represents 4% of all new cancer diagnoses in women yet the five-year survival is 42.6%. This represents a case-to-fatality ratio which is twice that of breast cancer, the most commonly diagnosed cancer in women (24.8% of all cancers in women)[3,4]. As such, ovarian cancer is the deadliest gynaecologic malignancy in the developed world. While treatments are improving, a lack of robust screening methods means that half of all new cases are diagnosed at a late stage (stage 3 or 4)[2].

Due to the overwhelming evidence that a significant proportion of 'ovarian' cancers originate in the fallopian tubes, as demonstrated in risk-reducing salpingo-oophrectomy specimens from women with *BRCA* mutations, in 2014 the Committee for Gynecology Oncology from the International Federation of Gynecology and Obstetrics (FIGO), updated the staging system to combine ovarian, peritoneal and fallopian tube cancer into one classification[5]. Whilst it is now accepted that potentially up to 60-80% of the most common ovarian cancer subtype, high grade serous cancer (HGSC), originates in the fimbrial end of the fallopian tube[6-9] , as the dissemination pattern and clinical behaviour of ovarian, peritoneal and fallopian tube are almost indistinguishable, the same management can be applied to all three diseases. These high-grade serous cancers represent over 70% of ovarian cancer deaths[10].

As a result of this relatively recent identification of the fallopian tube as the likely primary site of carcinogenesis, the presumed pathophysiology of epithelial ovarian cancer has shifted. Both ovarian epithelial and fallopian tube cancer are thought to arise due to *TP53* mutations, however, the historic theory of ovarian surface epithelial surface damage due incessant ovulation and repetitive damage has been superseded. The current concept describes a proinflammatory environment in the distal fallopian tube boosting the DNA damaging effects of the p53 signature, thereby encouraging the formation of serous tubal intraepithelial carcinomas (STIC) lesions, the precursor of fallopian tube carcinomas and HGSC[11].

1.2 RISK FACTORS FOR OVARIAN CANCER

Like many cancers, age is the biggest risk factor for ovarian cancer, largely due to the accumulation of DNA damage over time. Hereditary causes have been identified e.g., *BRCA1* and *BRCA2* mutations, however, due to the heterogenous nature of ovarian cancer, it has proved difficult to unequivocally describe non-genetic risk factors for ovarian cancer. Evidence has been demonstrated for obesity and tobacco smoking with further research required into other putative risk factors[12].

1.2.1 Genetic predisposition

It is estimated that 23% of ovarian, peritoneal, and fallopian tube cancers have a hereditary cause, of which the majority are pathogenic variants in the *BRCA1* or *BRCA2* gene[13]. Women who carry germline variants have a lifetime risk of ovarian, peritoneal, or tubal cancer of 20-50% for *BRCA1* and 10-20% with *BRCA2*, in comparison to 2% for the general population[6]. These cancers tend to develop at a younger age than the sporadic cases, with a median age of mid-40s in the case of *BRCA1* and after the age of 50 for *BRCA2*. Importantly, in over 40% of patients, they are the incident case in their family. Whilst it is well known that patients with *BRCA1* and *2* genetic mutations have an increased risk of ovarian and breast cancer (a lifetime risk of up to 80% of breast cancer), these patients also have an increased risk of several other cancers including pancreatic, melanoma and prostate cancer in men[14]. Several other hereditary ovarian cancers discovered via next generation sequencing are also associated with genetic anomalies, particularly DNA damage response pathway such as the mismatch repair genes in Lynch syndrome and *PALB2*[13].

1.2.2 Non-genetic risk factors

Eleven percent of ovarian cancer cases are preventable[15]. The International Agency for Research on Cancer (IARC) has classified several non-genetic exposures as increasing the risk of developing ovarian cancer including asbestos exposure, greater adult attained height, increased body fatness, oestrogen-only HRT (hormone replacement therapy), tobacco exposure, and perineal use of talc. They also designated oestrogen-progesterone contraception as protective factor against ovarian cancer[16]. Pelvic inflammatory disease has been suspected to play a significant role in pelvic malignancies for many years with a proposed causative mechanism revolving around a chronic infection-inflammation-neoplasia hypothesis. The oncogenic mechanisms behind most of these risk factors have not yet been fully elucidated.

1.2.2.1 Asbestos

The mechanism of asbestos-induced carcinogenesis in ovarian cancer is not well understood. Asbestos fibres have been found in ovarian tumours of women with work-related exposure in the Norwegian pulp and paper industry, and in women whose household contacts worked with asbestos. Two theories have been considered. One theory postulates that cell injury is induced by reactive oxygen species and hydroxyl radicals produced as a result of incomplete phagocytosis, where the macrophage is damaged because it is unable to digest the whole asbestos fibre due to its length. The second theory relates to mechanical irritation due to the asbestos fibres[17,18].

1.2.2.2 Greater adult attained height

Greater adult attained height has been implicated in several cancer conditions including malignant melanoma, breast cancer and colon cancer[19]. The underlying pathophysiology has not been fully elucidated and includes a basic principle that with more cells, there is an increased risk of genetic mutation and cellular change. Another consideration is for early life environment including nutrition and hormone levels. Genome-wide association studies (GWAS) have identified height specific gene loci that have also been implicated in cancer development thus suggesting that genetic and environmental interactions in early life may be related to cancer risk[20].

1.2.2.3 Endometriosis

Atypical endometriosis is the precursor lesion to both endometrioid and clear cell carcinomas of the ovary. Up to 42% of ovarian endometrioid cancers are associated with an ipsilateral ovarian or pelvic endometriosis[9]. A direct transition from ovarian atypical endometriosis to endometrioid or clear cell carcinoma has been described in 15-32% of cases[21]. In cases where an association between endometriosis and clear cell carcinoma has been noted, a more favourable prognosis is seen[22].

1.2.2.4 Pathogenic microbes, polymicrobial states and pelvic inflammatory disease (PID)

The past twenty years has seen a rapid progression in our understanding of our microbial cohabitants and their role in health and disease states. In the context of cancer development, eleven microbes or groups of microbes are now recognized as carcinogenic to humans[20]. Several of these carcinogenic microbes promote a chronic inflammatory state e.g., *Helicobacter pylori* and gastric cancer, and human papilloma virus (HPV) and cervical cancer. Shu et al., first made a case for an association between PID and ovarian cancer in 1989 when they noted that although statistically not significant, women with ovarian cancer were three times more likely to have a history of PID than controls (OR 3.0, 95%CI 0.3-30.2)[23]. Subsequently, several further studies have investigated the association with varying results (Table 1.1). Recent meta-analyses by both Zhou et al in 2017 and Piao et al in 2020 reported a positive association between PID and ovarian cancer (RR 1.24, 95%CI 1.06-1.44, HR 1.18, 95%CI 1.13-1.22 respectively) [3,24,25]. Furthermore, Piao et al., found positive associations between both invasive and borderline tumours (HR 1.28, 95%CI 1.20-1.30, HR1.28, 95%CI 1.19-1.37 respectively). Conversely, evidence suggests that tubal ligation reduces the risk of ovarian cancer. Both Cibula et al., 2011 and Wang et al., 2016 found a decreased risk of epithelial ovarian cancer following tubal ligation (RR0.66, 95%CI 0.60-0.73, OR0.70, 95%CI 0.60-0.81 respectively) despite using different data sets indicating that removing the passage from the endometrial cavity to the ovaries might mitigate the risk of ovarian oncogenesis[26,27]. Whether this reduction reflects a decrease in inflammation related to infection, retrograde menstruation or other unknown mechanisms remains to be explored

Table 1.1 Details of studies investigating the presence of microbes and ovarian cancer, and PID and ovarian cancer.

STUDY	STUDY CHARACTERISTICS	SUMMARY OF FINDINGS
Specific bacteria		
Chan et al., 1996[28]	27 samples of archived FFPE confirmed malignant ovarian tissue were examined for *Mycoplasma* using PCR-ELISA	16/27 (59.3%) of tissue samples were positive
Idahl et al., 2011[29]	291 samples: 209 benign, 16 BOT, 54 EOC, 12 other malignancies 271 matched controls <u>Plasma antibodies:</u> MIF of *C. trachomatis* IgG and IgA. *C. pneumoniae* IgG and chlamydial HSP60 IgG LAMP-EIA of *M. genitalium* IgG antibodies	No associations of antibodies with EOC Chlamydial hear shock protein 60 (HSP60-1) IgG were associated with type 2 OC in women with plasma obtained >1 year pre diagnosis *M. genitalium* IgG found in borderline tumours more than matched controls and benign conditions
Dadashi et al., 2016	Nested PCR of MOMP genes 23srrRNA *C. trachomatis* and 16srRNA *M. genitalium* of FFPE tissue OC:62, controls: 62	14/62 (22.5%) malignant samples were positive for *C. trachomatis* 0/124 positive for *M. genitalium*
Paradowska et al., 2019[30]	27: EOC (20 HGSC, 7 other OC, 4 metastatic), controls: 8 Ovarian/ fallopian tube tissue samples from unilateral side Nested PCR and qRT- PCR for CMV, EBV and HSV-1 PCR for HPV DNA <u>Plasma antibodies:</u> Anti-CMV IgG and IgM antibodies	CMV and HPV DNA detected in 70% and 74% of cancerous ovarian tissue, EOC vs benign P≤0.01 2/3 EOC coinfection CMV and HPV DNA CMV and HPV DNA also present in fallopian tube samples
Idahl et al., 2020[31]	EPIC study: 791 OC, 1669 matched controls Serum antibodies against *C. trachomatis*, *Mycoplasma genitalium*, HSV-2 and HPV16, 18 and 45	40% of study population seropositive for ≤STI. *C. trachomatis* antibodies associated with MC (RR2.30)

	Multiplex fluorescent bead-based serology	Chlamydia HSP60-1 associated with EOC (RR1.36), serous (RR1.44) HSV-2 associated with endometrioid (RR2.35)
Microbiome		
Shanmughapriya et al, 2012[32]	Fresh tissue samples of normal (9), benign (6) and neoplastic ovaries (24) DNA extraction by SDS-proteinase K-phenol chloroform method PCR detection of HPV, and nested PCR detection of C. trachomatis and CMV	15/39 (40%) HPV 6 (6/24 inv, 6/6 borderline, 3/9 normal) 24/39 (61.5%) *C. trachomatis* (21/24 inv, 3/6 borderline, 0/9 normal) 15/39 (40%) CMV (12/24 inv, 3/6 borderline, 0/9 normal)
Banerjee et al, 2017[33]	OC tissue: 99 Matched tissue (non cancerous tissue from ipsilateral or contralateral ovary): 20 Unmatched tissue (ovarian tissue from prophylactic surgery in *BRCA* positive women): 20 PathoChip Array, probe capture and NGS (Illumina MiSeq) for genomes of all known viruses, bacteria, parasites and fungi from FFPE tissue.	Viruses: 23% of those detected were tumorigenic and were in >50% of the cancer samples. Retroviridae highest hybridization. HPV 16, 18 in OC samples. Average hybridization higher in controls than in OC samples. Specific molecular signatures HPV 16/18 only detected in the cancer samples. Bacteria: Proteobacteria (52%), Firmicutes (22%), Bacteroidetes, Chlamydia Spirochaetes all in OC. None shared between all types of samples. Much more in OC vs controls. Shewanella in 91% of OC. Fungal: Rhizomucor, Rhodotorula, Alternaria, Geotrichum, Cladosporium in 95% of OC. Geotrichum only one shared.
Zhou et al, 2019[34]	Fresh HGSOC tissue: 71	Discovery:

	Fresh benign fallopian tube tissue: 66 Environmental controls: 12 16s rRNA sequencing in the discovery phase (25 OC, 25 controls) qPCR in the validation phase (46 OC, 41 controls) Transcriptome sequencing (RNA-seq) on signalling pathways on 6 OC and 7 controls	Decreased bacterial diversity in the OC tissue vs control group (Shannon P=0.0215), lower richness in the control group OC: Proteobacteria (upregulated), Firmicutes (downregulated), Bacteroidetes, Actinobacteria Validation: Significant differences in Proteobacteria and Fircmicutes between OC and controls (P<0.05). Actinobacter increased, Lactococcus decreased, Actinobater_lwoffi markedly enriched in OC Signalling pathways: Inflammation-associated signaling pathways activated in OC incl. NF-kappa B, Cytokine cytokine receptor interaction and Chemokine signalling pathway. TLR4/5, NLRP1, NOD1 reduced in OC, TLR6, MYD88 and NLRP3 increased
Nene et al, 2019[35]	Case-control study with two sets; ovarian cancer set and *BRCA* set EOC: Cases: 176, healthy:115, benign: 69 *BRCA*: *BRCA* +ve: 109, healthy (wild type): 97, benign (wild type): 14 16S rRNA gene sequencing of cervicovaginal swabs	EOC: ›50yoa CST O ‹50yoa EOC CST O ++vs ‹50yoa controls *BRCA*: ‹50yoa CST O vs ‹50yoa controls, ++ if 1st degree relative The younger the patient, the more strongly the CST O if OC or *BRCA* +ve

	CST collapsed into CST L (>50% *lactobacillus*), or CST O (<50% *lactobacillus*)	
Wang et al, 2020[36]	OC: 6, benign controls: 10 Ovarian tissue samples Immunohistochemistry staining using antibacterial LPS antibody to confirm presence of bacteria in ovarian tissue 16S rRNA sequencing to compare the taxonomic profile of the tissue	Bacterial lipopolysaccharide (LPS) present in both ovarian cancerous and non- cancerous tissue Different bacterial compositions in cancerous and non-cancerous ovarian tissue Gene functions associated with potentially pathogenic and oxidative-stress tolerant phenotypes enriched in cancerous ovarian tissue
Brewster et al, 2022[37]	10 women; 25 samples both OC and controls (don't specify the breakdown) Tissue samples of proximal fallopian tube, fimbriae and ovaries 16S rRNA sequencing	Presence of bacteria suggested in the UGT Different microbiome in each anatomical area; significant difference from fallopian tube to ovary, proximal fallopian tube different from fimbriae Different bacterial species enriched in ovaries and fallopian tubes of women with ovarian cancer
Qin et al, 2022[38]	OC: 35 (27 SC, 2 MC, 3 CC, 3 EC), benign controls: 30 11/35 OC did not have tubal ligation; 13/35 ligated Plat. resistant: 8; Plat. sensitive: 19 Swabs from cervix, uterine cavity, fallopian tubes and ovaries Metagenomic shotgun sequencing	Bacteria >99% of all microbes in ovaries OC distinct microbiota signature. Proteobacteria highest abundance Bacterial migration from the cervix noted. Higher α- diversity in ovaries than cervix, uterus, fallopian tubes

		No differences in bacterial on OC surface between patients with platinum- sensitive and platinum- resistant patients
PID		Case-control studies
Shu et al, 1989[23]	Set in Hong Kong, data gathered via interview	172 cases, 172 controls; OR 3.0, 95%CI 0.3-30.2
Risch et al, 1995[39]	Set in Canada, data gathered via interview	450 cases, 564 controls; OR 1.53, 95%CI 1.10-2.13[a], OR 1.88, 95%CI 1.13-3.12[b]
Parazzini et al, 1996[40]	Set in Italy, data gathered via questionnaire/interview	971 cases, 2758 controls; RR 0.7, 95%CI 0.4-1.3
Ness et al, 2000[41]	Set in the USA, data gathered via interview	616 cases, 1367 controls; OR 1.3, 95%CI0.6-2.5
Merritt et al, 2008[42]	Set in Australia, data gathered via questionnaire	1576 cases, 1509 controls; OR1.15, 95%CI 0.85-1.57
Lin et al, 2011[43]	Set in Taiwan, data gathered via database	67936 cases, 135872 controls; HR 1.92, 95%CI 1.27-2.92, HR 2.46, 05%CI 1.48-4.09[c]
Rasmussen et al, 2013[44]	Set in Denmark, data gathered via interview	554 cases, 1564 controls, OR 0.83, 95% CI 0.65-1.05
Rasmussen et al, 2017[45] *cohort	Set in Denmark, data gathered from nationwide Danish registries, population-based cohort study	1318929 population; 81281 cases of ovarian cancer, 5356 cases of PID HR 1.05, 95%CI 0.92-1.20 overall, HR 1.19, 95%CI 1.00-1.41 serous

Abbreviations: BOT, borderline ovarian tumours; *BRCA*, breast cancer susceptibility gene; CMV, cytomegalovirus; CST L/O, community state type lactobacillus/high diversity; DNA, deoxyribonucleic acid; EOC, epithelial ovarian cancer; FFPE, fresh frozen paraffin embedded; HGSOC, high grade serous ovarian cancer; HPV, human papilloma virus; IgA, immunoglobulin A; IgG, immunoglobulin A; inv, invasive; LAMP-EIA, lipid-associated membrane protein- enzyme immunoassay; MIF, microimmunofluorescence; MOMP, mitochondrial outer membrane permeabilisation; MYD, myeloid differentiation primary response; NLRP, nucleotide-binding oligomerization domain, Leucin rich Repeat and pyrin domain containing; NOD, nucleotide-binding oligomerisation domain-like receptor; OC, ovarian cancer; PCR, polymerase chain reaction; PCR-ELISA, polymerase chain reaction enzyme-linked immunosorbent assay; qPCR, quantatative polymerase chain reaction; rRNA, ribosomal ribonucleic acid;seq, sequencing; SDS-proteinase, sodium dodecyl sulfate proteinase; TLR, toll like receptor; yoa, years of age; +ve, positive.

[a] One episode of PID

[b] ≥Two episodes of PID

[c] ≥ Five episodes of PID

1.3 OVARIAN CANCER PREVENTION

1.3.1 Risk reducing bilateral salpingo-oophrectomy (RRBSO): primary prevention of ovarian cancer

The most lethal subtype of ovarian cancer, high grade serous (HGSC), accounts for up to 70% of sporadic ovarian cancers, and is strongly associated with *BRCA* variants. As a result of the current evidence demonstrating that HGSC likely originates in the fallopian tube, an opportunity for primary prevention in the general population has arisen in the form of opportunistic bilateral salpingectomy (BS) i.e., to remove both of the fallopian tubes whilst retaining the ovaries. Potential occasions to offer the procedure include at the time of hysterectomy or by removing the fallopian tubes in place of tubal ligation. Interestingly, retrospective data has suggested that those who have undergone BS due to inflammatory conditions such as hydrosalpinx or pelvic inflammatory disease experience a decreased risk of ovarian cancer [46]. Currently this procedure is not commonly performed in the UK with trials, such as the PROTECTOR study, ongoing to establish the efficacy and potential risks versus the current gold standard of bilateral salpingo-oophrectomy (BSO).

For those with an elevated risk due to *BRCA1/2* variant, elective BSO is the current gold standard of care. It reduces the risk of ovarian or fallopian tube cancers, primarily in those with *BRCA 2*, by 80% and has also been shown to decrease the risk of breast cancer[47,48]. Those with other genetic variants e.g., *RAD51C*, *RAD51D* have also been advised to undergo RRBSO to reduce their risk of high grade serous cancer. It has been demonstrated that opportunistic salpingectomy alone has minimal surgical risks perioperatively, has an unknown correlation with earlier menopause, and is cost-effective[49-52]. Comparatively, whilst those undergoing RRBSO also benefit from a reduction in cancer-risk, they are vulnerable to the side-effects of sudden surgical menopause, often without recourse for treatment due to associations of HRT and breast cancer[53].

1.3.2 Screening for ovarian cancer: secondary prevention of ovarian cancer

Over half of all ovarian cancer cases present with stage III or IV disease, both of which have a poor five year survival (27% for stage III and 13% for stage IV)[2]. Despite efforts over many years, effective screening methods for serous cancer have been elusive. All trials have included monitoring CA125 levels and/or transvaginal pelvic ultrasound scans, however, no combination has hitherto proved to be effective in reducing mortality due to ovarian cancer[54-57].

Carbohydrate antigen 125, CA125, is a glycoprotein biomarker that can be used in the diagnostic and monitoring pathways for ovarian cancer. There are, however, a broad variety of pathological and

physiological causes for it to rise (Figure 1.1). Traditionally, most assays consider <35mmol/L a normal level. CA125 is expressed by a wide range of tissues which have derived from embryonic coelomic epithelia including, for example, the endometrium, peritoneum, pleura and pericardium[58]. Both mechanical and inflammatory stress are the biggest stimulators of CA125 synthesis by mesothelial cells. Mechanical stress includes fluid congestion. Inflammatory stress includes the release of cell products e.g., tumour necrosis factors and interleukins[58].

CA125 has been demonstrated to be raised in several conditions of fluid overload in both benign and malignant conditions such as pulmonary oedema, liver cirrhosis, tumours of the digestive tract, along with serous cancers of the ovaries and fallopian tubes. Interestingly, a correlation was identified between volume of ascites and CA125, suggesting that the CA125 is not synthesised by the tumour specifically, and thus is not a tumour marker as such[59,60]. CA125 has also been demonstrated to be raised in conditions not related to serosal effusion such as pelvic inflammatory disease[61], endometriosis[62], lower respiratory tract infections and well-controlled congestive heart failure[63], suggesting a response to inflammatory stress.

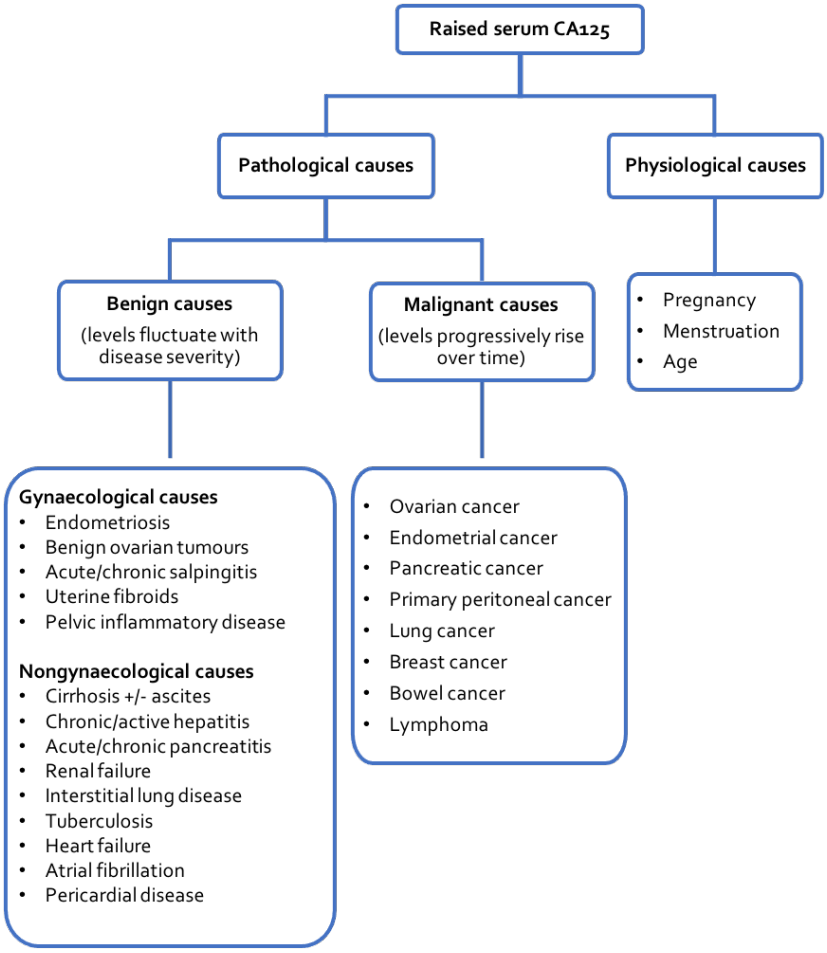


Figure 1.1 Summary of causes of a raised CA125.

Serum CA125 is used in several ways in the diagnosis and management of ovarian cancer. It is often raised in epithelial ovarian cancers, but less frequently in non-epithelial tumours[64]. It is elevated in only 50% women with stage 1 cancer, thus supporting the theory that it is not produced by the tumour itself[65]. The serum level may be used in conjunction with transvaginal ultrasound scanning of the pelvis when making a diagnosis. It is recommended as a first line investigation in patients with symptoms of ovarian cancer as per current national guidelines and is also used to calculate risk of malignancy in those with ovarian cysts via the risk of malignancy index (RMI)[66]. Following diagnosis of ovarian cancer, it is used to monitor response to treatment and identify recurrence.

Transvaginal pelvic ultrasound scanning is the gold standard for evaluating adnexal cysts. Ovarian masses are classified as 'simple' or 'complex'. Approximately 25% of ovarian masses cannot be classified on ultrasonagraphic qualities alone[67]. National guidelines from the Royal College of Obstetrics and Gynaecology (RCOG), advise that features of simple cysts include: thin-walled, anechoic, round or oval shaped, posterior acoustic enhancement and absence of septations or nodules. Complex masses have one or more of complete septation or multilocular nature, solid nodules or papillary projections. Should the mass extend beyond the scope of the transvaginal scan, a transabdominal ultrasound scan should be used as an adjunct. The International Ovarian Tumour Analysis group (IOTA), categorise ovarian masses into benign, malignant, or inconclusive based on ultrasonographic features alone (Table 1.2). When the IOTA rules are applied, of the 75% that can be classified, the test has 90% sensitivity and 95% specificity[68].

Table 1.2 Ultrasonographic features of benign and malignant ovarian masses as per the IOTA group. If the ovarian mass has at least one of the benign features with no malignant features it can be considered to be benign, and vice versa.

Benign	Malignant
Unilocular	Irregular solid tumour
Smooth multilocular tumour <10cm	Irregular multilocular- solid mass >10cm
Solid component <7mm in diameter	≥4 papillary structures
Presence of acoustic shadows	Ascites
No detectable Doppler signal	High Doppler signal (colour score 4)

Menon et al from the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS) group recruited 202,638 women to their randomized controlled trial assessing if population screening can reduce deaths from ovarian cancer[57]. Following a median follow up of 16.3 years, a significant reduction in mortality was not demonstrated in either of the two screening groups versus the control group; multimodal (repeated CA125 serum levels and transvaginal ultrasound scan (TVS)), TVS only, and no

screening. Some promising findings were noted in the MMS group, including a 39.2% increase in incidence of stage I and stage II disease, and 10.2% decrease in incidence of stage III and IV disease.

1.4 HISTOPATHOLOGY OF OVARIAN CANCER

Defining ovarian masses has proven to be complex due to the myriad of tissue types and cellular abnormalities that can occur. Their classification and histological examination is therefore appropriately complex and requires specialist input. Tumours can range from benign to borderline through to malignant neoplasms. Traditionally, tumours have been classified according to the microscopic cellular morphology of ovarian tissue. As a result of advances in molecular identification studies related to ovarian carcinomas however, the importance of incorporating both morphological and molecular information to accurately identify ovarian carcinomas is becoming a necessity[69].

Masses can initially be divided by their predominant histological type i.e., epithelial, non-epithelial or metastatic. Both epithelial and non- epithelial masses can be further divided into benign, borderline, and malignant based on their invasiveness and nuclear atypia. This division is vital as it reflects the cellular behaviour.

1.4.1 Epithelial tumours

A significant factor in the failure to establish an effective screening programme or tool for ovarian cancer is the heterogenous nature of the condition[8,57]. It has been demonstrated many times over that ovarian cancer is not a single disease, but a heterogenous collection of distinct carcinomas that can be classified based on different morphological and molecular genetic features[70]. Adjustments to classification systems are continuously ongoing due to new results from next generation profiling which encompasses high throughput, unbiased sequencing of whole genomes and transcriptomes of ovarian tissue, and is often combined with immunohistochemical testing to provide comprehensive results on all available molecular alterations[71].

The FIGO Cancer Report 2021 endorses the World Health Organisation's (WHO) histologic typing of epithelial ovarian tumours[6,9]. These malignant neoplasms, which account for the vast majority of tumours, comprise several subtypes including serous, mucinous, endometrioid, clear cell, Brenner, undifferentiated, mixed and a high grade serous carcinoma where the primary origin i.e., ovary or fallopian tube, is unable to be determined[5].

In 2010, Kurman and Shih initially proposed a dualistic model that categorised ovarian cancer into two distinct groups based on clinicopathological features and behaviours (Table 1.3)[1,69]. Type I tumours are

generally indolent in nature, confined to the ovary at presentation, genetically stable and low grade. They typically have well-defined and distinctly different morphological differences. There is often a clear lineage graduating from a benign cystic neoplasm through to an intermediate, borderline, tumour before progressing to cancer. In contrast, type II tumours have more overlap in their diagnosis due to less distinct morphological patterns. Type II tumours evolve rapidly, are aggressive and often present at an advanced stage with metastatic disease. This concept was revisited in 2016 when advances in molecular pathogenesis enable further subclassification and the emerging evidence related to the fallopian tube origins of HGSC[69]. Kurman's classification model has been subject to scrutiny however, with critics querying whether despite comprising only one-third of all ovarian carcinomas, low- grade serous carcinomas (LGSC), low-grade endometrioid (LGE), clear cell carcinomas (CCC) and mucinous carcinomas (MC) should be grouped together as 'Type I' tumours, due to their dissimilarity to high-grade serous carcinomas (HGSC), when in fact they are also clinically, morphologically and molecularly different from one another[72,73]. More commonly, epithelial ovarian cancer is divided into five subtypes which account for over 95% of all cases: HGSC, EC, CCC, MC and LGSC.

Table 1.3 Classification of ovarian tumours by Kurman and Shih[1]. Categorisation into two groups based on grade, aggressiveness, and morphological features.

Type I	Type II
Low- grade serous	High-grade serous
Low-grade endometrioid	High-grade endometrioid
Clear cell	Undifferentiated carcinomas
Mucinous	

1.4.1.1 High- grade serous carcinomas

High-grade serous carcinomas are the most common ovarian carcinomas (70%) with most patients presenting at an advanced stage, FIGO stage 3 or 4 (Table 1.4), (80%) with bilateral lesions (>90%) and a poor 5-year survival. Two histological types of HGSCs are discussed including classic-type and those with SET features (solid, endometrioid-like, and transitional)[6,73]. Classic-type tumours typically display increased architectural complexity, papillary and solid growth with slit-like glandular lumens showing foci of necrosis and exfoliation[70]. The tumour cells are of intermediate size, with scattered bizarre mononuclear giant cells with prominent nucleoli which have a wide-ranging variation in size[70,72]. They have a high mitotic activity. SET- type tumours consist of a solid mass of cells similar to endometrioid or transition cell carcinomas[73]. They have a greater number of tumour-infiltrating lymphocytes and a higher mitotic index than their classic-type counterparts.

1.4.1.2 Endometrioid carcinoma

Endometrioid carcinomas (ECs) account for 10-15% of all ovarian carcinomas[9,73]. They occur predominantly around perimenopause, are usually found at an early stage (Stage 1a) and subsequently have a more favourable prognosis. EC tumours are bilateral in 28% of cases (Stage 1b) and are associated with endometrial carcinoma in 15-20% of cases[9,73].

Most ECs are low-grade adenocarcinomas closely resembling uterine endometrial carcinomas[73]. They are well differentiated, although moderately and poorly differentiated carcinomas are occasionally found. Squamous differentiation occurs in 50% of cases[73]. Microscopically, foci of squamous metaplasia, not seen in (pure) variants of other ovarian carcinomas, are evident. The glands are crowded and fused but the glycocalyx is smooth and the epithelium is generally non-exfoliative. The nuclei are stratified and occupy the full thickness of the epithelium[70]. High-grade ECs are morphologically indistinguishable from HGSCs and gene expression profiling would suggest the high-grade EC is not a distinct tumour type[69,72,73].

1.4.1.3 Mucinous carcinomas

Mucinous tumours account for 3-4% of all primary ovarian carcinomas. They are composed of gastrointestinal-type cells containing intra-cytoplasmic mucin[9]. Typically, they are found in women in their 40's. They present early, usually confined to the ovary and advanced stage primary mucinous carcinoma is rare. Metastatic tumours from the gastric pylorus, intestine or endocervix can occur which may cloud diagnosis; however, most of these tumours show gastric or upper intestinal differentiation[73]. Primary MCs are usually large (>13cm) and unilateral, versus metastatic masses which are commonly smaller, bilateral and may be associated with pseudomyxoma peritonei[73]. They are usually well differentiated with moderate and poorly differentiated tumours relatively uncommon[69]. MCs are notably heterogenous with a continuum of architectural and cytological atypia that ranges from benign to borderline through to overtly carcinomatous areas. Microscopically, invasive carcinoma is characterised by two different patterns of invasion, both of which may be present within a single tumour. The expansile invasion or confluent glandular pattern is identified by closely packed malignant glands with little intervening stroma or invasion[9,73]. A cribriform pattern may also be evident. The second pattern, of destructive stromal invasion, is less common and is recognised by irregular glands, nests and single cells with malignant cytological features infiltrating stroma[9]. These have a high mitotic activity and abnormal mitotic figures are often present.

1.4.1.4 Clear cell carcinomas

Approximately 10% of ovarian carcinomas are clear cell carcinomas. CCCs are often diagnosed at Stage 1 or 2 with unilateral tumours[73]. There is a strong association with endometriosis, approximately 50-

70%, and the prognosis is more favourable in such cases. CCCs frequently contain a borderline component which is usually of an adenofibromatous nature. Histological diagnosis is made on the following features: clear cells with eccentric, rounded, bulbous nuclei, multiple complex papillae, densely hyaline basement membrane material expanding the cores of the papillae and hyaline bodies which are present in approximately 25% of cases[73]. Another useful histological finding is zonal tumour necrosis without associated stromal reaction which is often decisive during frozen section examination. In contrast to HGSCs, CCCs do not have *BRCA* abnormalities, chromosomal instability or complex karyotypes. TP53 mutations are also usually absent. The most common mutations noted include *ARID1A* (46-57%), *PIK3CA* activating mutation (40%) and loss of *PTEN* expression (8.3%)[9,73]. It has also been associated with Lynch syndrome[9].

1.4.1.5 Low-grade serous carcinoma

Low-grade serous carcinomas are uncommon accounting for less than 5% of all serous carcinomas[9,73]. They present at a younger age than HGSC, but also often at an advanced stage. Clinically, patients may present similarly to those with HGCS with bilateral masses and symptomatic of ascites however the tumour burden is often less. Prognosis of LGSC confined to the ovary and managed surgically is excellent, with five-year survival of advanced stage disease 87% and 10 year survival 50%. LGSCs do not respond well to traditional platinum- based chemotherapy. Microscopically, they demonstrate micro- or macropapillae surrounded by an unlined clear space, single cells with uniform small nuclei and irregularly shaped small nests of cells haphazardly infiltrating stroma with psammamoma bodies frequently noted. A serous borderline tumour (SBT) component is commonly noted. Necrosis is almost uniformly absent. Uniformity of nuclei, with limited nuclear pleomorphism is the main distinguishing feature between LGSC and HGSC. *KRAS* and *BRAF* mutations are seen in 50-60% of cases and relative to HGSC, the LGSC genome is relatively stable[9].

1.4.2 Fallopian tube and primary peritoneal cancers

As previously discussed, ovarian cancer was originally considered to be a cancer arising solely from the ovary, however it is now recognised as arising predominantly from the fallopian tubes, with lesions also arising from the pelvic peritoneum. This was discovered during sectioning of specimens removed during RRBSO in women with *BRCA* mutations and other risk factors for developing ovarian cancer, when small non-invasive and invasive carcinomas were discovered much more frequently in fallopian tubes than in ovaries[74,75]. The SEE-FIM pathology protocol (Sectioning and Extensively Examining the FIMbriated end) encouraged standardisation of examination of these specimens and led to reproducible identification of early serous carcinomas in the distal end of the fallopian tube[8]. The non-invasive lesions are designated as serous tubal intraepithelial carcinoma (STIC) lesions. Similar to

HGSC, they show *TP53* mutation and p53 protein expression and a high level of genomic instability.

A further 20% of ovarian cancer cases have complete obliteration of the fallopian tubes which may be the primary site[9]. Further proof is illustrated by identical *TP53* mutations in the STIC and the concurrent disseminated disease, with comparison of telomere length indicating the STIC lesion predated the ovarian tumour[9]. Ultimately, the investigation and management of these high grade pelvic serous cancers is the same.

1.5 STAGING AND GRADING OF OVARIAN CANCER

Pelvic cancer should be staged surgically, with computed tomography (CT) scan providing useful additional information. Occasionally in advanced disease, neoadjuvant chemotherapy is indicated. In such cases, histological or cytological confirmation of the diagnosis is required prior to starting treatment. The updated 2014 FIGO criteria may be applied to assign a stage (Table 1.4).

Table 1.4 2014 FIGO criteria for staging cancers of the ovaries, fallopian tubes and peritoneum.

FIGO Staging		
Stage I:		Tumour confined to ovaries or fallopian tube(s)
	1A:	Tumour limited to 1 ovary (capsule intact) or fallopian tube; no tumour on ovarian or fallopian tube surface; no malignant cells in the ascites or peritoneal washings
	1B:	Tumour limited to both ovaries (capsules intact) or fallopian tubes; no tumour on ovarian or fallopian tube surface; no malignant cells in the ascites or peritoneal washings
	1C:	Tumour limited to 1 or both ovaries or fallopian tubes, with any of the following:
		1C1: Surgical spill
		1C2: Capsule ruptured before surgery or tumour on ovarian or fallopian tube surface
		1C3: Malignant cells in the ascites or peritoneal washings
Stage II:		Tumour involves 1 or both ovaries or fallopian tubes with pelvic extension (below pelvic brim) or peritoneal cancer
	IIA:	Extension and/or implants on uterus and/or fallopian tubes and/or ovaries
	IIB:	Extension to other pelvic intraperitoneal tissues
Stage III:		Tumour involves 1 or both ovaries or fallopian tubes, or peritoneal cancer, with cytologically or histologically confirmed spread to the peritoneum outside the pelvis and/or metastasis to the retroperitoneal lymph nodes
	IIIA1:	Positive retroperitoneal lymph nodes only (cytologically or histologically proven):
		IIIA1 (i): Metastasis up to 10 mm in greatest dimension IIIA1 (ii): Metastasis more than 10 mm in greatest dimension
	IIIA2:	Microscopic extrapelvic (above the pelvic brim) peritoneal involvement with or without positive retroperitoneal lymph nodes

	IIIB	Macroscopic peritoneal metastasis beyond the pelvis up to 2 cm in greatest dimension, with or without metastasis to the retroperitoneal lymph nodes
	IIIC	Macroscopic peritoneal metastasis beyond the pelvis more than 2 cm in greatest dimension, with or without metastasis to the retroperitoneal lymph nodes (includes extension of tumour to capsule of liver and spleen without parenchymal involvement of either organ)
Stage IV:		Distant metastasis excluding peritoneal metastases
	IVA:	Pleural effusion with positive cytology
	IVB:	Parenchymal metastases and metastases to extra-abdominal organs (including inguinal lymph nodes and lymph nodes outside of the abdominal cavity)

Epithelial tumours can also be subclassified by histologic grading which is correlated to prognosis[9]. Two grading systems are in use; non-serous carcinomas are graded in an identical fashion as to the uterus, based on architecture with a one-step upgrade if there is nuclear atypia:

- GX: Grade cannot be assessed
- G1: Well differentiated
- G2: Moderately differentiated
- G3: Poorly differentiated

Serous carcinomas are divided into low and high grade. High grade tumours include both classic and those with SET (solid, pseudo-endometrioid, and transitional) features frequently have *TP53* gene mutations. Most moderately differentiated serous carcinomas carry mutations in *TP53* and are combined with high-grade tumours. Low-grade tumours often have borderline or atypical proliferative features and have mutations in *BRAF* and *KRAS*[6].

1.6 PATHOGENESIS OF OVARIAN CANCER

1.6.1 Molecular pathogenesis of ovarian cancer

Most HGSC express aberrant p53 levels (protein resulting from *TP53* gene), WT1 and p16. Estrogen receptor (ER) is expressed in approximately two-thirds of cases of HGSC and whilst it is also expressed in low grade serous ovarian carcinomas and endometrioid carcinomas, it is almost always negative in clear cell carcinomas and mucinous carcinomas[72]. As previous discussed, evidence has demonstrated that most HGSC derive from STIC lesions in the fallopian tubes. Small foci of strongly p53- immunoreactive cells were isolated from otherwise largely histologically normal fallopian tube epithelium, specifically in the distal portion of the tube, and have been designated 'p53 signatures'.

These p53 signatures subsequently develop into STIC lesions and finally invasive serous carcinoma. Those with *BRCA*1 or *BRCA*2 have a 17-44% risk of developing ovarian cancer by age 70, primarily HGSC. 5-10% of *BRCA*-positive asymptomatic women have been found to have early HGSC and 80% have associated STIC lesions when the histological specimens removed during prophylactic BSO have been examined.[6]. Despite these findings, not all HGSC display STIC lesions including those with SET morphology. Subsequently, HGSC may be further subdivided into those who are younger with *BRCA*, without STIC and a more favourable outcome, and those who are older, without *BRCA*, with STIC and a poorer outcome.

1.6.2 Proinflammatory conditions and ovarian cancer

Both endometriosis and PID are proinflammatory conditions with robust epidemiological evidence supporting a link with ovarian cancer[25,45,76-79]. Endometriotic cysts have elevated levels of iron which promotes oxidative stress and may stimulate genetic mutations. Patients with endometriosis have been found to have lowered levels of immune-suppressing activated T-cells in the endometrioma and endometrium, with T-cell depleted mouse models demonstrating increased levels of inflammatory cytokines and angiogenic factors[80]. Reactive oxidative species levels have also been noted to be higher in the blood and ovarian follicular fluid of women with endometriosis[81].

Chronic infections are associated with an increased number of cell divisions, which boosts the risk of mutations by increasing the potential for DNA replication errors during DNA repair. Inflammation releases many tumour-promoting mediators e.g., growth factors, cytokines, chemokines and free radicals which can produce a number of genetic and epigenetic alterations.[82-84] These alterations may lead to neoplasia and carcinogenesis by affecting pathways necessary for cellular homeostasis. For example, in the lower genital tract, *C. trachomatis* is thought to fulfil this role as a co-factor in cervical cancer development as it induces cervicitis which may predispose women to HPV infection through disruption of epithelial integrity.[85] In the upper genital tract, *C. trachomatis* is suspected to contribute to PID by causing a similar, chronic inflammatory environment.

1.7 TREATMENT OF OVARIAN CANCER

Prognosis of ovarian cancer is based on three main parameters including disease stage at diagnosis, histological type and grade and maximum diameter of residual disease after cytoreductive surgery. Both surgical and medical management in combination is the mainstay of treatment.

Surgical staging via a laparotomy with a systematic assessment of the peritoneal cavity including cytological and histological sampling throughout is the optimal procedure.

1.7.1 Cytoreductive surgery

1.7.1.1 Primary debulking surgery (PDS)

Full cytoreductive surgery should be offered in the first instance to patients with maximal effort to leave no residual disease. This comprises a primary laparotomy with total abdominal hysterectomy, bilateral salpingo-oophrectomy, omentectomy and frequently, bowel resection, with partial or complete multivisceral resection and removal of clinically suspicious pelvic and para-aortic lymph nodes.

1.7.1.2 Interval debulking surgery with neoadjuvant chemotherapy

For patients who are not fit for upfront debulking surgery, once Stage IIIC and IV disease is confirmed cytologically, weekly or three weekly cycles of chemotherapy are offered for three cycles before reassessing disease progression. If no progression is noted, with consideration of other factors such as patient performance status etc, interval debulking surgery is offered. The EORTC and CHORUS trials of neoadjuvant chemotherapy and interval debulking surgery demonstrated equivalent survival with less morbidity than PDS in selected patients[86-88].

1.7.2 Adjuvant chemotherapy

Patients with early-stage low-grade epithelial ovarian cancer i.e., Stage IA and Stage IB, who have optimal surgery gain no additional benefit from adjuvant chemotherapy. This is similarly the case for carcinoma *in situ* of the fallopian tube. Patients with high grade disease, or Stage IC and further, who undergo PDS are offered adjuvant three-weekly chemotherapy treatments for six cycles following surgery. Hyperthermic intraperitoneal chemotherapy is increasingly being adopted in the UK as another option in patients with advanced disease who have had optimal debulking surgery. However, clinical trials are still ongoing and it has not formally been introduced as part of national clinical guidelines.

1.8 THE HUMAN MICROBIOME

The human microbiome describes the entire habitat of a given site, including all micro-organisms, their genomes and the surrounding environmental conditions[89]. It has been the subject of intense and exponential investigation since the advent of metagenomics through both high-throughput analytical platforms and increased computational technologies[89].

The Human Microbiome Project (HMP), launched in 2007, aimed to use these metagenomic techniques to comprehensively characterize the human microbiome and analyse its role in human

health and disease[90]. The HMP comprised two phases; the first characterized the microbial communities in five different sites in the 'healthy' human body longitudinally and examined the roles of these microbes in human health and disease states. It also generated and made publicly available the vast amount of data that was collected to facilitate and support further research into this field by other groups. The second phase, Integrative Human Microbiome Project (iHMP), is focusing on examining three specific microbiome-related conditions, pregnancy and preterm birth, inflammatory bowel disease and type 2 diabetes. Longitudinal, collaborative studies are ongoing which are aimed at creating new multi-omic tools to examine the mechanisms linking these conditions, microbial activity and dysbiosis[91].

This sequential progression of HMP ventures from describing the microbial community structure, to investigating the microbial function related to disease states reflects the shift in microbiome research trends. It is recognized that clinical translation of microbial research lies in identifying the mechanisms and evolutionary relationships between dysbiotic communities and human disease states. Whilst the commonly quoted figure of 1:10 somatic to microbial cells has been debunked (it is closer to 1:1), the genome of the bacterial species in the human body has been underplayed and contains closer to 100 to 150 times the number of human genes[92,93]. With this knowledge, further multi-omic approaches to examine microbial metabolite and gene expression profiles have been developed including metagenomics, metabolomics and metatranscriptomics (study of microbiota genome and genes, analytical approaches used to examine the metabolite profile in a single strain and analysis of the expressed RNAs by high-throughput sequencing of corresponding meta-cDNAs, respectively) to facilitate further study[94,95]. Other large microbiome consortia such as the European Commission's Metagenomics of the Human Intestinal Tract project (MetaHIT), the Canadian Microbiome Initiative 1 and 2, the Irish Metagenomics of the Elderly programme (ElderMet) and the Japanese Human Metagenome consortium are adding to the global community effort to better understand this complex relationship[96-99].

1.8.1 The microbiome of the female genital tract

1.8.1.1 The microbiome of the lower genital tract

1.8.1.1.1 The cervicovaginal niche

Ravel et al., characterised the vaginal microbiome in 2011, describing five cohorts based on the relative abundance of bacterial composition[100]. These cohorts were referred to as community state types (CSTs) and have been the platform upon which further vaginal microbiome research has been based. CSTs I, II, III and V are characterised by enrichment of *Lactobacillus crispatus*, *L. gasseri*, *L. iners* and *L. jensenii* respectively and tend to have low bacterial species diversity and richness. Conversely, CST IV

is dominated by anaerobic bacteria frequently associated with bacterial vaginosis including *Gardnerella*, *Megasphera*, *Sneathia* and *Prevotella* species. Gajer et al., described the vaginal microbiome community further in 2012 when they executed a longitudinal study which illustrated the dynamic fluctuations of the vaginal microbiome composition[101].

Dysbiotic cervicovaginal microbiomes have been noted in cervical neoplasia, cervical, endometrial and ovarian cancers[35,102-104]. Mitra et al., noted that vaginal dysbiosis, characterized by high diversity and low levels of protective Lactobacillus species, was associated with progressive CIN with several bacterial species significantly correlating with high-grade disease, in particular *Sneathia sanguinegens*[104]. Walther-Antonio et al., demonstrated an over-representation of *Atopobium vaginae* and *Porphyromonas sp.* in the lower genital tract of patients with endometrial cancer[105]. A CST O, representing <50% abundance of protective Lactobacillus species, in the cervicovaginal niche, was found to be correlated with women aged 50 years and older with ovarian cancer and with women younger than 50 years with *BRCA1* mutation without ovarian cancer[35].

1.8.1.2 The microbiome of the upper genital tract

The upper genital tract is notably different from the lower genital tract in two main ways; it is much more diverse than the lower genital tract with all sites demonstrating several different bacterial phyla and has a significantly lower biomass than the lower genital tract, shown to be 1- to 4- orders of magnitude lower than the vagina and cervix[106]. Despite several studies in recent years, there is no consensus on a 'core' composition of the endometrial, fallopian tube or ovarian microbiotas.

1.8.1.2.1 The endometrium

In most studies examining benign conditions, *Lactobacillus* has been noted as the most abundant genus with several other genera also identified as colonisers including *Bacteroides*, *Gardnerella*, *Staphylococcus* etc[105-109]. Conversely, a Lactobacillus-deplete composition has been noted in the endometrial samples of patients with endometrial cancer, with an enrichment of *Dialister*, *Anaerococcus*, *Prevotella*, *Porphyromas* and *Peptococcus* species[105,110]

1.8.1.2.2 The fallopian tubes and ovaries

There is scant evidence assessing the presence or composition of the microbial landscape of both the fallopian tubes and ovaries. Sampling these areas is hindered both by difficulty in accessing these sites and extremely low biomass, which makes sampling prone to contamination. Of the evidence that is available, the upper genital tract appears to be as equally as diverse as the endometrium, with

Firmicutes, Proteobacteria, Actinobacteria, Bacteroidetes all being noted in the fallopian tubes and ovaries[33,34,106,111]

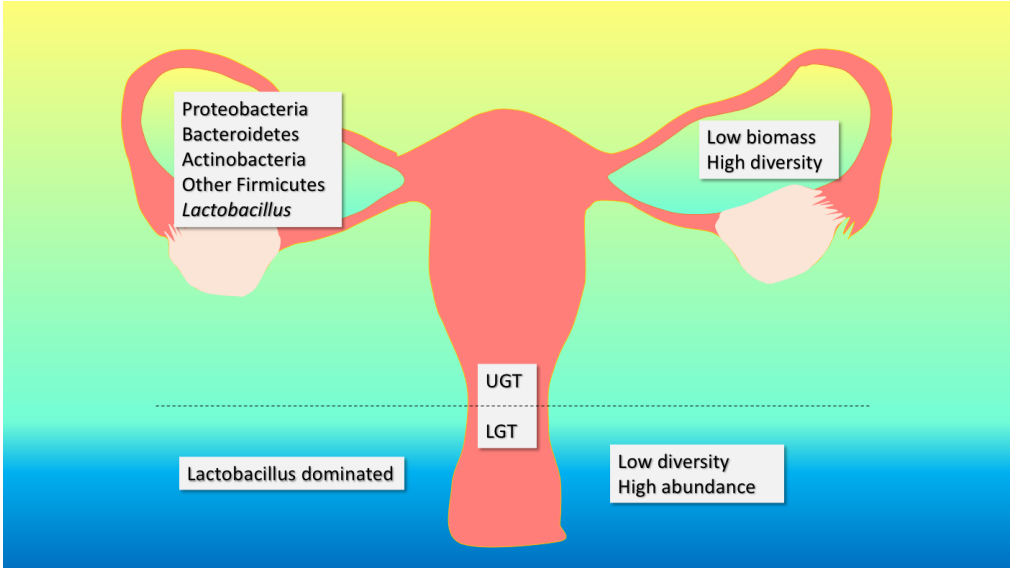


Figure 1.2 Common phyla in the female genital tract.

1.9 HYPOTHESIS AND AIMS

1.9.1 Rationale

Ovarian cancer is the deadliest gynaecologic malignancy in the developed world with few modifiable risk factors. Modifiable risk factors that have been identified in the literature e.g., obesity, diabetes and PID, are related to a proinflammatory environment. While treatments are improving, a lack of robust screening methods means that almost six in 10 cases are diagnosed at a late stage [2].

The pathophysiology of ovarian cancer is still not fully understood due to the heterogenous microenvironment of ovarian tumours. It is well established that 15-20% of all high-grade serous ovarian carcinoma (HGSC) are caused by germline and somatic mutations in the *BRCA* genes, with a further 30-35% caused by other homologous recombination repair genes [8].

It has recently been accepted that a substantial proportion of ovarian cancers develop from the fallopian tube and serous tubal intra-epithelial carcinomas (STIC) lesions have been located in the fallopian tube of approximately 50-60% of HGSCs [7,8].

Further work is required to understand the initiating events in the development of ovarian cancer at this site.

Since the inception of next generation sequencing 15 years ago, the microbiome and its role in health and disease has been under intense scrutiny. The current thinking is that the microbiome can be viewed as a newly recognised immunological organ, which works harmoniously with our bodily functions in healthy states but can be disrupted by a variety of factors not limited to diet, surgery and medications thereby contributing to disease. A significant proportion of work is studying inflammation and the microbiome, a process which has been implicated in initiating ovarian oncogenesis[112,113].

There are clear and well-defined microbiomes in several body locations e.g., alimentary canal, lungs, skin, vagina, however many sites previously considered sterile are currently being investigated for the presence of microbiota. Small studies have demonstrated that the upper genital tract displays a distinct microbiome [106,108]. Walther-António et al. demonstrated that there is a discrete microbiome in endometrial cancer relative to benign conditions [105]. Nené et al. demonstrated that both women with ovarian cancer or women who were high risk for ovarian cancer (*BRCA1* mutation) were significantly more likely to have a dysbiotic vaginal microbiome [35]. Banerjee et al also described specific microbes which were distinct to ovarian cancer tissue [33].

A plausible theory linking the microbiome and malignant transformation in the upper genital tract is inflammation, already a proven core process in the development of many other cancers e.g., colorectal, pancreatic, SCC, melanoma. Inflammation has been shown to be present both in the tissue surrounding neoplastic tumours and also within the tumours themselves[83]. Hallmarks of chronic or 'smouldering' inflammation in the tumour microenvironment which promote tumour growth include the presence of chemokines, cytokines and prostaglandins, tissue remodelling, angiogenesis and tissue repair[84]. These inflammatory-like processes aid in the proliferation of malignant cells, promotes angiogenesis and metastasis, subverts adaptive immune responses and alters responses to hormones and chemotherapeutic agents[84]. Chronic inflammation is not a novel concept in ovarian oncogenesis and was previously considered the underlying mechanism in the theory of 'incessant ovulation'. This describes a process of repetitive injury to the ovarian surface epithelium resulting in chronic inflammation leading to DNA damage produced by oxidative stress. This mechanism has been supercded by the fallopian tubes as the primary site of serous ovarian cancer, with resulting focus on causes of inflammation to the fallopian tubes e.g., PID and other dysbiotic states. Several epidemiological studies have described a correlation between pelvic inflammatory disease and ovarian cancer, and a protective effect of bilateral salpingectomy[40,43,114-116].

A potential mechanism linking microbes, inflammation and ovarian oncogenesis involves the suppression of p53, the protein resulting from the tumour suppressor gene, *TP53*. Mutated *TP53* is found in >96% of HGSC tissues. Wild-type *TP53* is known as the 'guardian of the genome' and is

responsible for cell apotosis, cell senesence and ensuring that defective cells are halted before overgrowth occurs. *TP53* mutations are believed to occur early in ovarian carcinogenesis, even in precancerous lesions e.g., STIC lesions [117]. NF-κB is a protein complex involved in cellular responses to many stimuli e.g., bacteria and stress. NF-κB suppresses the actions of p53 and it has been found to be upregulated in the malignant ovarian tumour tissue[34,118].

Several proinflammatory factors have been proven to be correlated with an increased risk of ovarian cancer e.g., genital talcum powder use, endometriosis and PID. The discovery of microbes in the upper genital tract through NGS has allowed the development of a novel hypothesis involving a new proinflammatory factor, previously left unexamined due to the lack of appropriate tools. Microbes leading to a proinflammatory state has the potential for therapeutic manipulation in ways other proinflammatory risk factors do not. The local vaginal microbiome, through production of metabolites, is thought to modify host cell signalling and immune pathways that impact local homeostasis, such as increasing the risk of pelvic inflammatory disease when dysbiosis, a state of microbial imbalance, is present. If a high-risk microbiological pattern or species is determined, it could be used to identify those women most in need of treatment. It may be possible to manipulate the vaginal microbiome through microbial gene regulation, probiotic and prebiotic treatments in order to establish a healthy environment throughout the female genital tract, combat disease and improve health outcomes.

1.9.2 Hypothesis

Building upon the current knowledge base of the role of the microbiome and inflammation in gynaecological cancers, I hypothesise that the female upper reproductive tract has distinct bacterial signature that differs between women with ovarian cancer and benign disease. Furthermore, I hypothesise that the microbiota in the upper reproductive tract forms a continuum with the microbiota in the lower genital tract with the same bacterial species found in all anatomical sites.

1.9.3 Aims

To this end, the aims of my thesis are:

1. To conduct an umbrella review investigating the strength and validity of existing epidemiological evidence for associations between non-genetic risk factors and the risk of developing or dying from ovarian cancer.
2. To describe the composition of the microbiota of the female reproductive tract in women with both malignant ovarian disease and benign disease.
3.

a. To investigate if the microbiota of the lower female reproductive tract correlates with the microbiota of the upper female reproductive tract
b. To identify any microbes specific to ovarian cancer throughout the female reproductive tract, focusing specifically on the easily accessible lower reproductive tract.

CHAPTER TWO: MATERIALS AND METHODS

2.1 MATERIALS

Table 2.1 List of materials and reagents used.

Experimental process	Chemical/kit/buffer	Supplier
General	Ethanol	VWR
	Methanol	VWR
	Phosphate buffered saline (PBS)	In house
	Tris/Borate/EDTA (TBE)	In house
	Sodium Citrate	Sigma
	Sterile distilled water	Gibco
	0.2ml PCR tubes	Starlabs
	0.5ml Eppendorf tubes	Starlabs
	1.5ml Eppendorf tubes	Starlabs
	2ml Eppendorf tubes	Starlabs
	10µL sterile filter tip	Starlabs
	20µL sterile filter tip	Starlabs
	200µL sterile filter tip	Starlabs
	1000µL sterile filter tip	Starlabs
	96-well optical plate	Starlabs
Sample collection	BBL™ CultureSwab™ cont. liquid Amies	Becton Dickinson
	Speculum	Pelican
16s rRNA extraction	DNA Mini Kit	Qiagen
	Lysostaphin	Sigma
	Lysozyme	Sigma
	Mutanolysin	Sigma
	QiAMP *Cador* Pathogen Kit	Qiagen
	TE (10mM Tris, 1mM EDTA at pH7.4)	In house
	Triton	Sigma
	Tissue Lyser LT	Qiagen
	0.1mm zircona/silica beads	BioSpec
	5mL syringes	Greiner Bio-One
PCR	Agarose	Sigma
	AmpliTaq Gold DNA Polymerase	ThermoFisher Scientific
	BioRad Tetra 2 machine	BioRad
	Deoxynucleotide mix (dNTP) 10mM	ThermoFisher Scientific
	$MgCl_2$ 25 μ M	ThermoFisher Scientific
	PCR-grade sterile water	Sigma
	SYBR® Safe DNA gel stain	ThermoFisher Scientific
	6x Blue/Orange Loading Dye	Promega
		Invitrogen

	16s rRNA forward & reverse primers	
Software	Excel	Microsoft
	PRISM	GraphPad
	R	R Foundation for Statistical Computing
	Revman	Cochrane
	Statistical Analysis of Metagenomic Profiles (STAMP)	Oximetrics
	ClustVis	
	MicrobiomeAnalyst	https://biit.cs.ut.ee/clustvis/ www.microbiomeanalyst.ca

2.2 METHODS FOR THE UMBRELLA REVIEW

2.2.1 Literature search

PubMed, EMBASE and the Cochrane Database of Systematic Reviews were searched from inception to September 2019 for systematic reviews and meta-analyses of observational and interventional studies that explored the association between non-genetic risk factors and the risk of developing or dying from ovarian cancer (Table 2.2). A further manual examination of the citations of the retrieved eligible papers to identify additional publications that might have been missed during the initial search was also performed.

Table 2.2 Search algorithm for ovarian cancer umbrella review

PubMed and Embase	
Exposure terms	Hip circumference OR waist circumference OR waist to hip ratio OR weight OR height OR weight gain OR obesity OR obese OR BMI OR Body Mass index OR bariatric surgery OR insulin OR metformin OR antiglycaemics OR aspirin OR aspirin containing medication OR paracetamol OR painkillers OR Diabetes OR type 2 diabetes mellitus OR birth weight OR exercise OR physical activity OR radiation OR radiotherapy OR hormone replacement therapy OR HRT OR HPV OR human papillomavirus OR breast cancer OR cancer OR tamoxifen OR smoking OR smoker OR tobacco OR contraception OR OCP or oral contraceptives OR oral contraceptive pills OR ethnic origin OR ethnicity OR race OR PCOS OR polycystic ovarian syndrome OR family history OR familial OR menarche OR menopause OR premenopausal OR post menopausal OR parity OR pregnancy OR alcohol OR coffee OR caffeinated intake OR caffeine intake OR caffeine OR decaffeinated OR decaffeinated intake OR marital status OR married OR single OR divorced OR cohabiting OR living together OR partner OR partnered OR breastfeeding OR endometrial hyperplasia OR socioeconomic status OR

	poverty OR wealth OR education OR level of education OR educational level OR schooling OR past gynaecological history OR endometriosis OR fibroid* OR non-cancerous cyst* OR pelvic inflammatory disease OR PID
	AND
Primary outcome terms	ovarian cancer OR ovarian carcinoma OR ovarian neoplasm OR ovarian tumor* OR ovarian tumour* OR adenocarcinoma OR cystadenoma* OR cystadenocarcinoma* OR adenoma* OR endometrioid carcinoma* OR ovarian maligna* OR borderline
	AND
Review terms	systematic review OR meta-analysis OR metaanalysis
Cochrane Database of Systematic reviews	
Exposure terms	Hip circumference OR waist circumference OR waist to hip ratio OR weight OR height OR weight gain OR obesity OR obese OR BMI OR Body Mass index OR bariatric surgery OR insulin OR metformin OR antiglycaemics OR aspirin OR aspirin containing medication OR paracetamol OR painkillers OR Diabetes OR type 2 diabetes mellitus OR birth weight OR exercise OR physical activity OR radiation OR radiotherapy OR hormone replacement therapy OR HRT OR HPV OR human papillomavirus OR breast cancer OR cancer OR tamoxifen OR smoking OR smoker OR tobacco OR contraception OR OCP or oral contraceptives OR oral contraceptive pills OR ethnic origin OR ethnicity OR race OR PCOS OR polycystic ovarian syndrome OR family history OR familial OR menarche OR menopause OR premenopausal OR post menopausal OR parity OR pregnancy OR alcohol OR coffee OR caffeinated intake OR caffeine intake OR caffeine OR decaffeinated OR decaffeinated intake OR marital status OR married OR single OR divorced OR cohabiting OR living together OR partner OR partnered OR breastfeeding OR endometrial hyperplasia OR socioeconomic status OR poverty OR wealth OR education OR level of education OR educational level OR schooling OR past gynaecological history OR endometriosis OR fibroid* OR non-cancerous cyst* OR pelvic inflammatory disease OR PID
	AND
Primary outcome terms	ovarian cancer OR ovarian carcinoma OR ovarian neoplasm OR ovarian tumor* OR ovarian tumour* OR adenocarcinoma OR cystadenoma* OR cystadenocarcinoma* OR adenoma* OR endometrioid carcinoma* OR ovarian maligna* OR borderline

2.2.2 Eligibility criteria

Systematic reviews or meta-analyses of observational studies in humans that assessed non-genetic risk factors and ovarian cancer incidence and mortality were included. Meta-analyses of randomised

controlled trials investigating the impact of interventions e.g., dietary, exercise, or mixed interventions on ovarian cancer incidence and mortality were also included.

Meta-analyses that reported benign ovarian pathology as an outcome (such as benign ovarian masses), and studies assessing genetic factors as the exposure of interest, as well as studies investigating the prognostic ability of factors on survival of patients previously diagnosed with ovarian cancer were excluded[119]. Meta-analyses that did not present comprehensive study specific data (number of cases of ovarian cancer or deaths, number of study population or person years, relative risks with corresponding 95% confidence intervals or standard error) were excluded. These missing data were not retrievable from the original studies or after contacting the authors of the meta-analyses.

In the event of multiple meta-analysis reporting on the same exposure-outcome associations, the paper with the largest number of studies to preclude duplication of the original studies was selected. Manual searches were performed to ensure primary studies were not included in more than one systematic review that formed the evidence base. The concordance between the included and duplicate meta-analyses was explored in a sensitivity analysis.

2.2.3 Evaluation of the strength of evidence

Each identified risk factor and ovarian cancer association was graded into four groups; strong, highly suggestive, suggestive and weak evidence using a series of criteria described previously[120-125]. These included: summary random-effects P-value; P-value of largest study in meta-analysis; number of cases; between-study heterogeneity; 95% prediction interval; small study effects; excess significance bias.

2.2.4 Evaluation of the quality of included meta-analyses

The strength and quality of all included meta-analyses was assessed using the AMSTAR 2 tool, which uses 16 measures to appraise the methodological quality of systematic reviews into high, moderate, low or critically low quality[126].

2.2.5 Data Extraction

The name of the first author and the year of publication, the modifiable exposure and outcome studied, and the summary effect (relative risk, odds risk, hazard ratio, standard incidence ratio, standard mortality ratio), and 95% confidence intervals (CI) or standard error (SE) were extracted from each eligible meta-analysis. Similarly extraction of the first author and the publication year, epidemiological design (cohort, case control), number of cancer cases and controls in case- control studies or the

number of cases and total population or person years in cohort studies, maximally adjusted relative risk (odds ratio in case-control studies, risk ratio or hazard ratio in cohort studies and the 95% confidence intervals were extracted.

Meta-analyses were evaluated as they were originally presented; expansion of one meta-analysis with studies detected by another on the same topic was beyond the scope of this review.

2.2.6 Data Analysis

The primary analysis focused on cohort studies as these studies provide most robust evidence among observational studies. Sensitivity analyses were conducted after including case-control studies.

2.2.6.1 Assessment of summary effect, heterogeneity and prediction intervals

For each exposure, the summary effect and 95% confidence interval was calculated using both fixed and inverse variance weighted random effects methods[127]. The Cochran Q test and the I^2 metric of inconsistency including its 95% confidence intervals were both implemented to assess for inter-study heterogeneity[128-130]. The I^2 metric ranges between 0% and 100% and could reflect either genuine diversity within the studies, chance or bias. A further assessment of inter-study heterogeneity was performed by calculating the 95% prediction intervals (PI) for the summary random effect estimates. The PI enables direct comparison with future clinically relevant effect estimates as it provides a range within which the effect estimate is predicted to fall[131,132].

2.2.6.2 Assessment of small study effects

To evaluate whether the smaller studies in a meta-analysis were creating an exaggerated risk estimate relative to the larger studies, Egger's test at $p<0.10$[133] with the the random-effects summary estimate being further away from the null-value compared to the point estimate of the largest study in a meta-analysis (i.e. the study with the smallest standard error) was used. Small study effects can occur due to publication or reporting bias, true heterogeneity or chance[134].

2.2.6.3 Evaluation of excess statistical significant bias

The test for excess significance aims to evaluate whether the observed (O) number of statistically significant results ($p<0.05$) in the studies included in a meta-analysis is too large compared to the expected (E) number[130]. The number of significant studies that were expected in each meta-analysis was calculated using the sum of the statistical power estimates for each included study using non-central *t*-distribution[130,135]. The effect of the largest study (i.e. has the smallest standard error) was used as the plausible effect size as the true effect size is not known. Sensitivity analyses were performed

using the summary fixed and random effect estimates as alternative plausible effect sizes. Excess significance for each individual meta-analysis was defined as two-sided $P<0.10$.

2.2.6.4 Evaluation of the quality of included meta-analyses

To date there is no quality assessment tool specifically designed for umbrella reviews. The AMSTAR 2 tool was used as a proxy to assess the methodological quality of the 73 included publications[126]. The World Cancer Research Fund Continuous Update Project (WCRF CUP) report was not included in this assessment, having already been subjected to extensive peer review processes[12].

The AMSTAR 2 tool places more weight on seven of the sixteen questions, with unsatisfactory answers to any of the seven being regarded as a critical weakness. The four grades range from 'high' to 'critically low' with 'high' having zero or one non-critical weakness in the study and 'critically low' having more than one critical flaw with or without non-critical weaknesses. The critical weaknesses detail an 'a priori' published protocol, a comprehensive literature search, a list of excluded studies, satisfactory risk of bias assessment for each included study, appropriate statistical analysis, adequate discussion regarding any bias detected and an appropriate investigation of publication bias. As the AMSTAR 2 tool aims to evaluate systematic reviews of randomised and non-randomised trials, some of the criteria were not applicable to umbrella reviews.

2.2.6.5 Evidence Grading

A description of the individual assessment criteria used to classify meta-analyses into evidence grades is presented in the table below:

Table 2.3 Details of criteria for evidence grading.

Assessment criteria / Evidence Grade	Summary random effect (*P*-value)	No. of cases	Heterogeneity (I^2)	95% prediction interval exc. null hypothesis	Small study effects	Excess significance	*P*-value of largest study
Strong	$<10^{-6}$	>1,000	<50%	Yes	None demonstrated	None demonstrated	x
Highly suggestive	$<10^{-6}$	>1,000	x	x	x	x	<0.05
Suggestive	$<10^{-3}$	>1,000	x	x	x	x	x
Weak	<0.05	x	x	x	x	x	x

2.2.6.6 Sensitivity analysis and credibility ceilings

A sensitivity tool called a credibility ceiling which is based on the assumption that a single observational study has a probability *c* (credibility ceiling), or maximum certainty (100-*c*)%, that the true effect size of the study is in a different direction from the one suggested by the point estimate was used[136]. The credibility ceiling at which the association was not significant was presented. The summary effect sizes and the inter-study heterogeneity were re-examined using credibility ceilings. By using this tool in sensitivity analysis, the aim was to account for possible methodological limitations of the observational studies, which can cause overestimation of the combined effect estimates[136,137]. For each meta-analysis and given a ceiling of c%, the likelihood ratio of the real effect size being in the direction indicated by the summary estimate for the corresponding unit increase (or level) of exposure was computed[136].

All statistical analyses were performed using Stata version 15 (College Station, TX) (StataCorp 2015), and all P values were two tailed[138].

There was no patient involvement.

2.3 METHODS FOR THE ANALYSIS OF THE FEMALE GENITAL TRACT MICROBIOTA

2.3.1 Study population

Non-pregnant women attending outpatient gynaecology oncology clinics at Imperial College NHS Healthcare Trust who were scheduled for a total hysterectomy and bilateral salpingo-oophrectomy were prospectively recruited. A further group of non-pregnant women scheduled for the same procedure were recruited on the day of the procedure from the gynaecology ward at Imperial College NHS Healthcare Trust.

Women were included irrespective of age, parity, menopause, ethnicity, smoking, contraception use, hygiene practices (e.g., douching), or previous gynaecological history. Data was collected specific to recent sexual activity (‹48 hours prior to swab collection) or recent antibiotic use (in the two weeks prior to sample collection). Exclusion criteria included history of HIV positive status, hepatitis B or C, autoimmune disorders or had a prior history of total hysterectomy or salpingo-oophrectomy.

2.3.2 Sample collection and processing

Samples were collected from patients once the patient was under general anaesthetic, before aseptic cleaning and standard pre-operative intravenous antibiotics were given. Cervicovaginal secretions were collected from the posterior vaginal fornix with a BBL™ CultureSwab™ containing liquid Amies (Becton Dickinson, Oxford, UK) using a sterile, disposable speculum (Pelican Feminine Healthcare, Cardiff, UK), without the lubricant. Swabs were stored on dry ice until all samples from the theatre list had been collected and were then immediately stored at -80°C. To facilitate completion of sample-taking from the other sites in the upper genital tract, the specimen comprising the uterus, fallopian tubes and ovaries was placed into a sterile bag following removal during the operation, stored initially at 4 °C before being transported to the pathology laboratory on the same day.

To ensure a sterile field for the specimen dissection in the pathology laboratory, a sterile drape, sterile gloves and sterile, single-use instruments were provided for the pathology registrar performing the dissection. Swab samples were collected from the endocervical junction, the fundus, the ostia of the fallopian tube and the external surface of the ovaries using a BBL™ CultureSwab™ containing liquid Amies (Becton Dickinson, Oxford, UK). Tissue samples were also taken by the pathology registrar from the same sites when the clinical situation allowed. A basic nutrient agar Petri dish was placed on the dissecting bench to capture any air-borne contaminants and was swabbed at the end of the dissection. All samples were stored at -80°C within 30 minutes.

A comprehensive questionnaire was used to obtain all relevant gynaecological, medical and surgical history (location). Menopausal status, type of contraception or hormone replacement therapy (HRT) use, previously or currently, and menstrual cycle phase (follicular or luteal) were documented.

2.3.3 Microbiome bacterial DNA extraction

Whole-genomic bacterial DNA was extracted from BBL™ CultureSwab™ samples using QiAmp Mini DNA kit (Qiagen, Venlo, Netherlands). Each swab to be processed was removed from the -80°C freezer and thawed on ice. Once the samples defrosted, they were vortexed vigorously to resuspend the cells in transport buffer (Amies Liquid Medium). The swab sponge was carefully transferred to a 5ml syringe where approximately 500µl of solution containing whole genomic bacterial and viral DNA was collected into a sterile DNAse- and RNAse-free 2ml centrifuge tube by compressing the swab between the plunger and syringe. The solution was centrifuged 5000 *x g* for 10 minutes to form a pellet. The Amies supernatant was removed and stored at -20 °C. The pellet was resuspended in a lysis buffer cocktail containing 50µl lysozyme (10mg/ml)(Sigma, Dorset, UK), 6µl mutanolysin (25,000U/ml)

(Sigma), 3µl lysostaphin (4,000 U/ml) (Sigma), 41µl TE_{50} (10mM Tris, 50mM EDTA at pH7.4), 30µl 12% Triton and 170µl phosphate buffered saline (PBS) per sample, at 37°C for one hour. The pellet was then resuspended by gentle vortexing and immediately underwent a further mechanical lysis step. Following this, 100mg of bleached and rinsed 0.1mm diameter zircona/silica beads were added to each sample and oscillated for one minute at 25Hz in a Tissue Lyser LT (Qiagen).

After brief centrifugation to collate the condensation and centralise the beads, the supernatant lysate was transferred to fresh 2ml centrifuge tubes and 200µl supernatant aspirated and applied to the QIAmp cador Pathogen Kit (Qiagen, Manchester, England) and used according to the manufacturer's protocol. Lastly, 100µl elution buffer (supplied with the kit) was added to give 100µl template DNA.

2.3.4 16S rRNA polymerase chain reaction (PCR)

Following the bacterial DNA extraction, a quality control step was performed where the V1-V2 hypervariable regions of 16SrRNA genes were amplified by polymerase chain reaction (PCR) using a forward and reverse fusion primer (Invitrogen, Carlsbad, CA USA) to confirm the presence of bacterial DNA prior to MiSeq sequencing[139]. For this, PCR mastermix was prepared as per Table 2.4, containing 2µl template DNA, to give a total of 50µl per reaction. Purified bacterial DNA was used as a positive control, whilst PCR-grade sterile water (Sigma) and an unused extracted swab were used as negative controls. PCR was run on a BioRad Tetra 2 machine (BioRad) using the conditions detailed in Table 2.5. Eight microliters of 6x Blue/Orange Loading Dye (Promega) was added to each PCR tube and the PCR products were run on a 2% agarose gel, containing SYBR® Safe DNA gel stain (ThermoFisher Scientific, Waltham, MA USA) with 1x Tris/Borate/EDTA (TBE) used as running buffer. The was viewed under ultraviolet (UV) light.

Table 2.4 Bacterial PCR Mastermix (Total 50µl per reaction)

Item	Volume per reaction
10 x AmpliTaq Gold Buffer (ThermoFisher Scientific)	5µl
$MgCl_2$ 25µM (ThermoFisher Scientific)	5µl
AmpliTaq Gold DNA Polymerase (ThermoFisher Scientific)	0.25µl
16s rRNA forward primer (Invitrogen) Sequence: 3' GCC TTG CCA GCC CGC TCA GTC AGA GTT TGA TCC TGG CTC AC	1µl
16s rRNA reverse primer (Invitrogen) Sequence: 5' GCC TCC CTC GCG CCA TCA GCA CTG CAT GCT GCC TCC CGT AGG AGT	1µl
Deoxynucleotide mix (dNTP) 10mM (ThermoFisher Scientific)	1µl
PCR-grade sterile water (Sigma)	34.75µl

Template DNA (from bacterial DNA extraction)	2µl

Table 2.5 PCR conditions for 16S rRNA primers

Step 1. Incubate at 95°C for 5 minutes
Step 2. Incubate at 95°C for 30 Seconds
Step 3. Incubate at 55°C for 30 Seconds
Step 4. Incubate at 72°C for 1 minute 30 Seconds
Step 5. Cycle to step 2 for 30 more times
Step 6. Incubate at 72°C for 7 minutes
Step 7. Incubate at 4°C till removed

2.3.5 Sequencing of bacterial 16S rRNA genes

Profiling of the bacterial component of the microbiota was performed using MiSeq platform (Illumina, San Diego, CA, USA) in the Department of Digestion, Metabolism & Reproduction (St Mary's Hospital, London, UK). The V1–V2 hypervariable regions of 16S rRNA genes were amplified for sequencing using the forward primer, 28F primer (5′-GAGTTTGATCNTGGCTCAG-3′) and the reverse primer, 388R primer (5′-TGCTGCCTCCCGTAGGAGT-3′). PCR amplification was performed in 25µl reaction volumes with Qiagen HotStar *Taq* master mix (Qiagen Inc, Valencia, California), 1µl of each 5uM primer, and 1µl of template on ABI Veriti thermocyclers (Applied Biosytems, Carlsbad, California) using the following thermal profile: 95°C for 5 min, 35 cycles of 95°C for 30 sec, 55°C for 40 sec, 72°C for 1 min, followed by one cycle of 72°C for 10 min and a 4°C hold. Amplicon libraries were cleaned and size selected using Agencourt AMPure XP (BeckmanCoulter, Indianapolis, Indiana) following the manufactures protocol. Pooled sequence libraries were quality checked and quantified using gel electrophoresis, Qubit 2.0 Fluorometer (Life Technologies) and quantitative PCR and loaded at 10pM on an Illumina MiSeq (Illumina, San Diego, California) using 2x300bp chemistry.

2.3.6 Processing of sequence data

The 16S rRNA gene sequence data was processed within the software package Mothur using the MiSeq SOP Pipeline[140]. Briefly, sequence reads were quality checked and singleton operational taxonomic units (OTUs) and OTUs with fewer than 10 reads in any sample were excluded. OTUs were defined using a cut off value of 97% and resulting data analysed using the *Vegan* package in R (R Foundation

for Statistical Computing) for assessment of microbial composition and diversity. OTU taxonomies (from Phylum to Genus level) were classified using the RDP MultiClassifier script to generate the RDP taxonomy[141]. The USEARCH algorithm combined with the cultures representatives from the RDP database were used to determine the species level taxonomies of the OTUs[142]. Alpha and beta diversity indices were calculated from these datasets with Mothur and using the *Vegan* package in R. Processing of 16S rRNA gene sequencing data in Mothur using the MiSeq SOP pipeline and assignment of OTU taxonomies were performed by Dr Ann Smith.

2.3.7 Statistical analyses

This study was designed as a descriptive, exploratory, hypothesis-generating experiment. Due to the nature of the study, a sample size calculation was not performed. This is in accordance with several other exploratory microbiome studies in similar areas[36-38,143]. Several different statistical platforms and packages were used to analyse statistical differences of microbiome data from different sites between the different populations. Statistical tests were selected depending on normality and the number of groups analysed. For both p-value and q-value, <0.05 was considered significant. The population was categorised into two groups of interest for the main analysis, women with benign disease and women with an ovarian malignancy. Differences in categorical clinical parameters between the two groups of interest were assessed primarily using Fisher's exact test for each of the listed characteristics (ethnicity, parity, menopausal status, HRT use, OCP use, history of PID and BMI) whilst a t-test was used for age.

A correlation analysis was performed in the Statistical Analysis of Metagenomic Profiles (STAMP) software package (Oxmetrics, Timberlake, London, UK) between samples taken from the endocervical junction and the fundus of the uterus at both species and genera level using Pearson's correlation coefficient[144].
MicrobiomeAnalyst (McGill University, Montreal, Canada), an online statistical analysis platform, was used to compare the microbiome data from samples retrieved from specimens removed laparoscopically and specimens removed transabdominally[145,146]. Univariate analysis was performed at genus level using Kruskall-Wallis tests.

Further univariate analysis was performed at species level to determine the presence of bacterial genetic data in the upper genital tract above background signature using Mann-Whitney and Kruskal-Wallis tests with multiple testing corrections applied using the Benjamini-Hochberg false discovery rate.

Microsoft Excel (Version 16.43, 2020) and Graph-Pad PRISM v9.3 (GraphPad Software, San Diego, CA) were used to produce the descriptive statistics of the 16S rRNA sequence read counts. The Kruskall-Wallis test was applied to most of the comparisons with t-test applied for the vaginal samples. Venn diagrams illustrating any bacterial species that were common throughout the genital tract of a patient were created using an online platform (https://bioinformatics.psb.ugent.be/webtools/Venn/).

Two methods of clustering were employed to determine if there were any groupings created by anatomical site at both species and genus level. Hierarchal clustering using Ward linkage was performed on the online statistical platform ClustVis for all sample sites. *Va*gina*L* community state typ*E* Nearest Centro*I*d classifier, or VALENCIA, an algorithm based on nearest-centroid classification specifically of vaginal samples was also implemented for the vaginal and cervical samples due to the elevated levels of bacterial crossover in the cervicovaginal niche[147].

The correct statistical test was utilized depending on whether the data was normally distributed or not, and the number of groups analysed (Figure 2.1). A p-value ‹0.05 was considered significant.

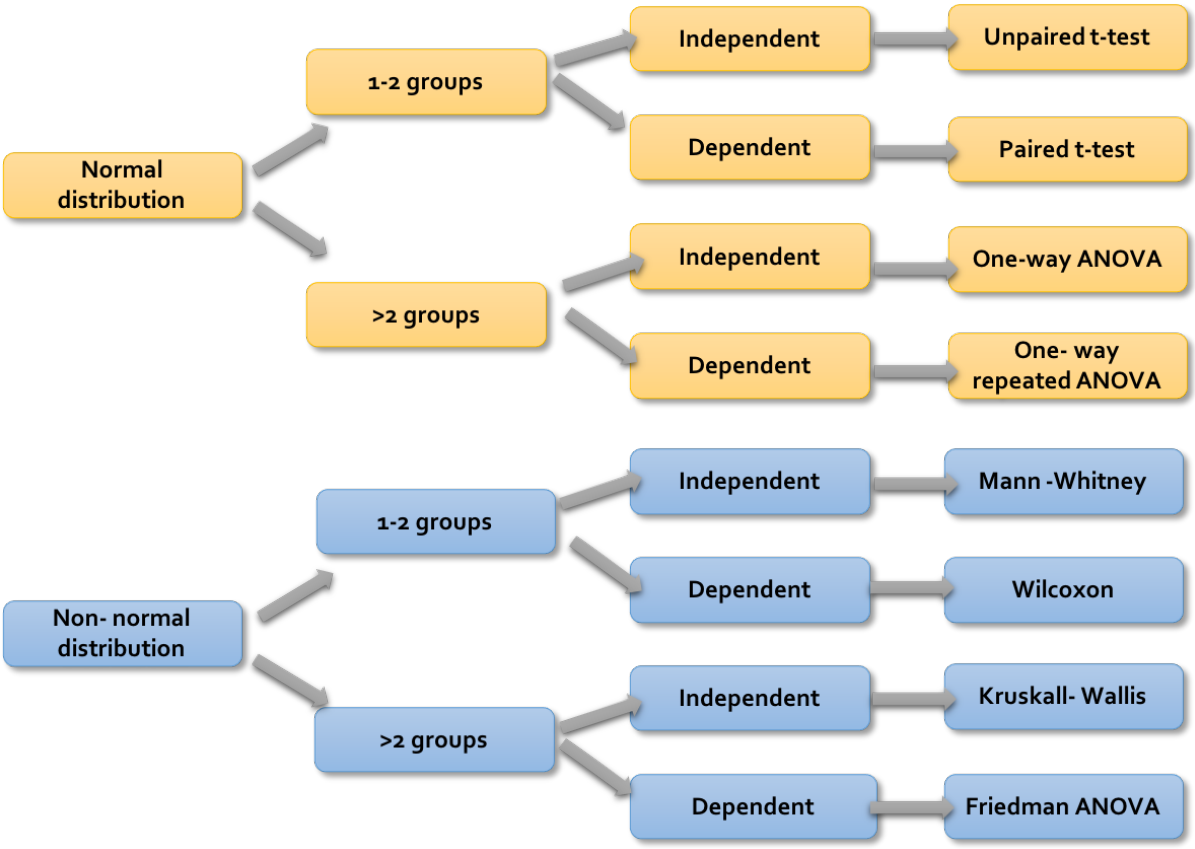


Figure 2.1 Description of selection of statistical analyses used in this thesis.

2.3.8 Contamination considerations in microbiome research

A known challenge in microbiome analysis of low biomass samples is minimising the risk of background, non-patient contaminants. In this study, samples from the upper genital tract were considered to be particularly prone to background contamination which would result in contaminant reads accounting for a larger proportion of the total read count than the samples with higher true bacterial biomasses (i.e., vagina and cervix). Several steps were taken to identify and remove any reads identified as possible background, non-patient contaminants.

2.3.8.1 Processes used to identify and manage background contamination in low biomass samples

A manual method of identifying and removing suspected contaminated reads was initially applied based on the concepts of frequency and prevalence. Frequency refers to the concept that contaminant DNA sequence reads will be present in equal and low amounts across all samples, representing a sizable proportion of low-concentration samples, versus true samples where DNA concentrations can vary widely. Prevalence infers that sequences from contaminants are present in higher proportions in negative controls than in true samples[148].

We also used the decontam package to identify and remove potential contaminants from the sequencing data. This is a recently developed standardized and validated method of assessing contaminated DNA sequence reads in microbial data. For this approach, two cut-offs where selected; $P<0.5$ was used for a more stringent cut off (potentially resulting in more false negatives) and $P<0.1$ less stringent (potentially resulting in more false positives).

CHAPTER THREE: RISK FACTORS FOR OVARIAN CANCER: AN UMBRELLA REVIEW OF THE LITERATURE

Content from this chapter has been published:

Whelan E, Kalliala I, Semertzidou A, Raglan O, Bowden S, Kechagias et al. Risk Factors for Ovarian Cancer: An Umbrella Review of the Literature. Cancers (Basel). 2022 May 30;14(11):2708. doi: 10.3390/cancers14112708.

3.1 ABSTRACT

3.1.1 Background

Several non-genetic factors have been associated with ovarian cancer incidence or mortality. Some of these associations may be flawed due to intrinsic biases. My objective was to evaluate the strength and validity of this evidence using an umbrella review methodology.

3.1.2 Materials and methods

Design: Umbrella review of meta-analyses.

Data sources: PubMed, EMBASE, Cochrane Database of Systematic Reviews, manual reference screening.

Eligibility criteria: Systematic reviews or meta-analyses that evaluated the association between non-genetic risk factors and ovarian cancer incidence and mortality.

Main Outcomes: The primary analysis focused on meta-analyses of cohort studies.

Data Synthesis: The evidence was graded into strong, highly suggestive, suggestive or weak based on statistical significance of the random-effects summary estimate of the largest study in a meta-analysis, number of cases, between-study heterogeneity, 95% prediction intervals, small study effects, and presence of excess significance bias.

3.1.3 Results

Two hundred and twelve meta-analyses, investigating 55 non-genetic risk factors for ovarian cancer were identified. Risk factors were grouped in eight categories: anthropometric indices, dietary intake, physical activity, pre-existing medical conditions, hormonal therapy use, biochemical markers, past gynaecological history and smoking. Of the 174 meta-analyses of cohort studies assessing 44 factors, six associations were graded with strong evidence. Greater height (relative risk (RR) per 10cm 1.16, 95% confidence interval (CI) 1.11 to 1.20), body mass index (RR ≥30kg/m^2 versus normal 1.27, 95% CI 1.17 to 1.38) and three exposures related to hormone replacement therapy use increased the risk of developing ovarian cancer. Use of oral contraceptive pill reduced the risk (RR 0.74, 95% CI 0.69 to 0.80).

3.1.4 Conclusions

Many risk factors have been associated with ovarian cancer, only six had strong evidence to support this association without bias. Refining the significance of genuine risk factors for the development of ovarian cancer may increase awareness in women at risk, aid prevention and early detection.

3.2 AIMS

The aims of this umbrella review are to:

1. Summarize the available evidence of meta- analyses and systematic reviews on non-genetic risk factors and developing or dying from ovarian cancer.

2. Assess the validity of the current epidemiological evidence on the many putative non-genetic risk factors for ovarian cancer and investigate the supposed most robust associations.

3.3 INTRODUCTION

Ovarian cancer is the most lethal gynaecological malignancy in high-income countries[149]. Worldwide, 434,184 new ovarian cancer diagnoses and 293,039 deaths are predicted to occur in 2040, an increase from 295,414 and 184,799 respectively in 2018[150]. In the United Kingdom (UK), over 7,400 cases are diagnosed every year and 4,100 die from the disease[2]. Although recent improvements in treatment with a combination of surgery, chemotherapy and more recently PARP inhibitors have led to an increase in 5-year survival for all stages from 21% in 1970 to 42% in 2017, the overall cure rates and prognosis remain poor[2,151]. Despite ovarian cancer causing only 2.5% of cancers in UK women, it accounts for 5% of all deaths[2]. Poor outcomes have been predominantly related to delays in diagnosis and absence of prevention, as nearly four out of five women are still diagnosed with advanced disease[2,151]. Concerted efforts explored the value of screening with disappointing results, whilst identification of women at high-risk may enhance prevention and early detection[152].

Although approximately 20% of the cases of ovarian cancer have been associated with genetic mutations in *BRCA1* and *BRCA2*, the majority of cases are sporadic. Increased body mass index (BMI) and hormone replacement therapy (HRT) have been proposed as major contributors along with smoking and occupational hazards e.g., asbestos exposure[15,151]. Conversely, use of the combined contraceptive pill has been suggested to reduce the risk of ovarian cancer[153]. Although many of these suggested associations could be true, some may well be biased. Residual confounding and selective reporting of strong positive results could result in false positive findings and overinflate the observed magnitudes of effect[154-156]. Umbrella reviews can systematically appraise the evidence presented in meta-analyses exploring associations between putative risk factors and related outcomes. Recent umbrella reviews have shown that despite claims of significant associations between several risk factors and outcomes, only a fraction were supported by robust evidence without hints of bias [123-125,157-160].

In this study, I present an umbrella review of systematic reviews and meta-analyses of observational studies to systematically appraise the evidence and validity of associations between non-genetic risk factors and the risk of developing or dying from ovarian cancer.

3.4 RESULTS

3.4.1 Characteristics of meta-analyses

From the literature search, 72 eligible publications were identified that included 212 meta-analyses of 1853 individual study estimates exploring associations between non-genetic risk factors and the risk of developing or dying from ovarian cancer (Figure 3.1)[12,17,24,27,119,153,161-229]. No meta-analyses of randomized controlled trials investigating the impact of interventions on ovarian cancer incidence and mortality were identified. Out of the 1853 study estimates, 946 (51%) were cohort studies, 795 (43%) case-control studies, 47 (2.5%) were nested case-control, seven (0.4%) were case cohort, one was a pooled analysis of cohort studies and seven (0.4%) were randomised controlled trials. There were two to 40 study estimates combined per meta-analyses with a median of six. The median number of cases and total population in each meta-analysis was 2,873 and 304,057, respectively. The lowest number of cases in a meta-analysis was 20 and the highest was 60,524, whereas the smallest total population or controls was 16 and the highest was 20,307,196. In 181 out of the 212 included meta-analyses, there were more than 1,000 cases of ovarian cancer.

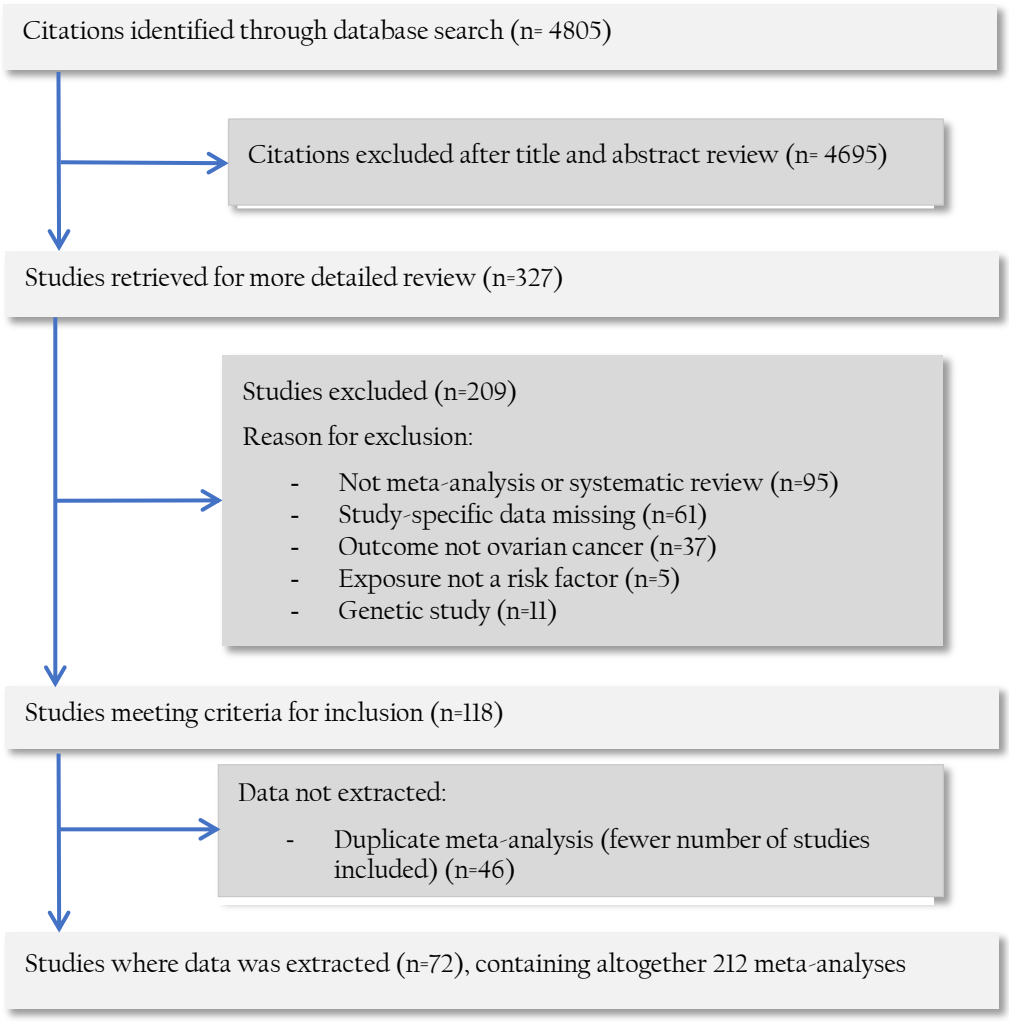


Figure 3.1 Flow diagram of selection of meta-analyses on risk factors for ovarian cancer.

A total of 55 risk factors were examined in the 212 meta-analyses, belonging broadly to eight categories; seven anthropometric indices, 17 dietary factors, two risk factors including physical activity and sedentary behaviour, three risk factors associated with pre-existing medical conditions or interventions, nine factors related to medication or hormonal therapy use, four biomarkers, 11 risk factors related to past gynaecological history and three carcinogens (Table 3.1).

Of the 212 meta-analyses, 174 included at least two cohort studies assessing 44 risk factors and were included in the main analysis. Eight of the 174 meta-analyses reported on ovarian cancer incidence and mortality, one meta-analysis reported on ovarian cancer mortality only, and the remaining meta-analysis reported on ovarian cancer incidence.

Table 3.1 Details of risk factors assessed.

Risk Factor Categories		
Anthropometric *n*=7	Hip circumference Waist circumference Waist to hip ratio Weight gain Weight BMI Height	
Dietary *n*=17	Dairy Type of dairy intake Fruit Meat Eggs Polyphenols Vegetables Alcohol Tea	Fish Fat Acrylamide Dietary inflammatory index Fibre Protein Coffee Dietary vitamins and minerals
Physical activity and sedentary behaviour *n*=2	Recreational activity Sedentary behaviour	
Past medical history *n*=3	SLE DM PCOS	
Past drug history *n*=9	Aspirin NSAIDS Acetaminophen HRT OCP	Statins Metformin Bisphosphonates Antidepressants
Biomarkers n=4	CRP IL6 TNFα TNFR2	
Past gynaecological history *n*=11	Number of terminations Age at menarche Tubal ligation Hysterectomy PID Breastfeeding	Multiple births Parity IVF Age at last birth Endometriosis
Carcinogens *n*=3	Talcum powder Smoking Asbestos	

Abbreviations: BMI, body mass index; CRP c- reactive protein; DM, diabetes mellitus; HRT, hormone replacement therapy; IL6, interleukin 6; IVF, in vitro fertilisation; NSAIDS, non-steroidal anti-inflammatory drugs; OCP, oral contraceptive pill; PCOS, polycystic ovarian syndrome; SLE, systemic lupus erythematosus; TNFα, tumour necrosis factor α; TNFR2, tumour necrosis factor receptor 2

3.4.2 Summary effect size

With p<0.05 as the threshold of statistical significance, the summary fixed and random effects estimates were significant in 39 (22%) and 40 (23%) out of the 174 meta-analyses of cohort studies, respectively (Table 3.2). At P<0.001, 22 (13%) and 17 (10%) meta-analyses produced significant summary results using the fixed and random effects model, respectively. At $P<10^{-6}$, summary fixed effects estimates were significant in 14 (8%) meta-analyses and summary random effects were significant in 10 (6%) meta-analyses. Of the 10 meta-analyses with strongly statistically significant summary random effect estimates, eight described an increased risk of ovarian cancer incidence or mortality for the following risk factors: obesity, height, hormone replacement therapy (HRT), endometriosis and asbestos exposure. Two meta-analyses including exposure to the oral contraceptive pill (OCP) and metformin demonstrated a decreased risk of ovarian cancer incidence.

Assessment of between-study heterogeneity, 95% prediction intervals, small study effects, and excess significance bias were performed only on those meta-analyses that demonstrated statistically significant results i.e., P<0.05, as previously performed by He and colleagues[230]. The meta-analyses with non-significant results are presented in Supplementary Tables S1 and S2.

The association of the largest study included in each meta-analysis was nominally statistically significant in 20 of the 40 statistically significant meta-analyses (50%), and the relative risks of the largest studies were more conservative than the summary random effects in 28 (70%) meta-analyses.

3.4.3 Heterogeneity between studies

Seven (18%) of the 40 meta-analyses presented significant evidence of between-study heterogeneity at P≤0.10 using the Cochran's Q test. Large heterogeneity (I^2 = 50-75%) was found in five (13%) meta-analyses and very large heterogeneity (I^2 >75%) in two (5%) meta-analyses for five risk factors including endometriosis, diabetes mellitus, HRT, in vitro fertilisation (IVF) and pelvic inflammatory disease (PID) (Table 3.3).

3.4.4 Small study effects

Four meta-analyses (10%) reported small study effects (Egger's test at P<0.10 with more conservative effects in the largest study of a meta-analysis compared to the summary random effects estimate) for endometriosis, HRT and BMI. (Table 3.3).

3.4.5 Excess significance

Five (12.5%) meta-analyses demonstrated evidence of excess significance bias using the largest study estimate as the plausible effect size (P<0.10). These included varying exposures, namely endometriosis, BMI, asbestos exposure, and breastfeeding. Using either the summary fixed or random effects estimates instead of the largest study estimate did not highlight any further meta-analyses with excess significance bias (Table 3.3).

3.4.6 Quality assessment

Using the AMSTAR 2 tool, all 72 publications included in this review were assessed. The WCRF CUP report was not included in this assessment, having already undergone rigorous peer review procedures[12]. 76% (55/72) of included meta-analyses stated 'a priori' published protocols and 89% (64/72) described a literature review. Only 4% (3/72) both listed excluded studies and justified their exclusion, whilst a further 35% (25/72) declared the excluded studies without explanation. Most meta-analyses provided satisfactory methods when assessing for risk of bias (78%, 56/72), when combining the findings (97%, 70/72), when accounting for risk of bias from individual studies in interpretation of the main results (81%, 58/72) and when investigating and discussing any small study bias (81%, 58/72). Of the remaining components of the AMSTAR 2, the majority of included meta-analyses fulfilled the criteria except for completing the study selection in duplicate (44%, 32/72) and assessing the potential impact of risk of bias of individual studies on the main results (33%, 24/72).

Thirty-five percent of all the included papers were graded as high (5/72, 7%) or moderate (20/72, 28%) quality, whilst 26% (19/72) and 39% (28/72) were graded as low or critically low (AMSTAR Table 3.5). Papers assessed with 'low' or 'critically low' quality failed to meet 'critical' criteria i.e., lack of protocol, literature search, description of excluded studies and risk of bias assessment.

One of the five included publications considered to be of high quality (AMSTAR Table 3.5), included one of the publications presenting strong evidence on the association between BMI and increased OC incidence[179,193,198,202,231]. Three publications which included the other five meta-analyses presenting strong evidence were deemed to be of low (height and increased OC incidence) or critically low quality (HRT and increased incidence of OC and OCP use and decreased incidence of OC)[153,161,169]. Grading of low or critically low quality was due to lack of description for the excluded studies[161] and lack of assessment for publication bias[153,169].

The three meta-analyses that received highly suggestive evidence had high quality (endometriosis), or moderate quality (ever vs. never use of estrogen only HRT and metformin use)[207,213,231].

Table 3.2: Description of 40 meta- analyses results investigating risk factors associated with ovarian cancer incidence or mortality –cohort studies.

Author, year	Exposure	Exposure contrast	N[α]	Sample size cases/ cohort	Summary relative risk (95% CI)			Fixed P-value[ε]	Random P-value[ϕ]	95% Prediction interval[γ]
					Fixed Effects[β]	Random Effects[χ]	Largest Study[δ]			
Anthropometric indices										
Pooralajal 2014	BMI	BMI ≥ 30kg/m^2 vs normal, PrMP	3	71/ 350211	1.57 (1.20-2.06)	1.57 (1.20-2.06)	1.56 (1.14-2.16)	9.7E-04	9.7E-04	0.27-9.02
Pooralajal 2014	BMI	BMI ≥ 30kg/m^2 vs normal	13	6947/20, 560, 388[ε]	1.27 (1.17-1.36)	1.27 (1.17-1.38)	1.27 (1.19-1.36)	2.8E-16	2.6E-8	1.09-1.47
Aune 2015	BMI	per 5kg/m2 increase	24	17734/ 16300000	1.06 (1.03-1.08)	1.07 (1.04-1.11)	0.97 (0.93-1.01)	2.9E-0.7	1.2E-04	0.96-1.21
Aune 2015	BMI	iya per 5kg/m2 increase	6	9452/ 11100000	1.12 (1.05-1.19)	1.12 (1.05-1.19)	1.16 (1.04-1.29)	3.8E-04	3.8E-04	1.03-1.23
Aune 2015	Weight	Per 5kg	4	1006/ 297350	1.03 (1.01-1.05)	1.03 (1.01-1.05)	1.02 (1.00-1.05)	8.0E-04	1.4E-03	0.98-1.08
Aune 2015	Height	Per 10cm	16	18663/ 13600000	1.14 (1.12-1.18)	1.16 (1.11-1.20)	1.14 (1.10-1.18)	1.17E-26	2.20E-13	1.06-1.26
Keum 2015	Weight gain	Per 5kg; never HRT (PoMP)	2	217/ 23984	1.13 (1.03-1.24)	1.13 (1.03-1.24)	1.16 (1.03-1.31)	8.5E-03	8.5E-03	NA
Dietary intake										
Song 2017	Calcium intake	HvL	5	1726/ 351192	0.86 (0.74-1.00)	0.86 (0.74-1.00)	0.86 (0.68-1.10)	4.00E-02	4.0E-02	0.67-1.09
Larsson 2006	Dairy	HvL, total dairy products	2	427/ 90001	1.66 (1.19-2.31)	1.66 (1.19-2.31)	1.61 (1.07-2.42)	2.8E-03	2.8E-03	NA
Larsson 2006	Dairy	HvL, skim/low fat	3	728/ 170327	1.35 (1.09-1.68)	1.35 (1.09-1.68)	1.32 (0.97-1.82)	5.7E-03	5.7E-03	0.35-5.43
Larsson 2006	Dairy	HvL, lactose	3	728/ 170327	1.47 (1.17-1.84)	1.47 (1.17-1.84)	1.48 (1.05-2.09)	7.8E-04	7.8E-04	0.34-6.29
Kolahdooz 2010	Meat	HvL, processed	3	1018/ 696100	1.26 (1.02-1.56)	1.26 (1.02-1.56)	1.23 (0.92-1.63)	3.5E-02	3.5E-02	0.31-5.07
Wallin 2011	Meat	Per 100g/wk increment; red and processed	21	6536/ 2140286	1.01 (1.00-1.04)	1.01 (1.00-1.04)	1.02 (0.98-1.06)	3.4E-02	3.4E-02	1.00-1.04
WCRF CUP 2013	Non starchy vegetables	Per 100g/day	6	2053/ 641079	0.95 (0.90-0.99)	0.94 (0.89-1.00)	1.00 (0.93-1.07)	2.8E-02	4.0E-02	0.82-1.08
Zhang 2018	Non herbal tea	HvL	3	734/ 164882	0.69 (0.52-0.93)	0.69 (0.52-0.93)	0.63 (0.40-0.99)	1.4E-02	1.4E-02	0.11-4.57
Butler 2011	Tea; black	HvL	5	1299/ 203998	0.72 (0.57-0.91)	0.73 (0.56-0.93)	0.63 (0.40-0.99)	5.4E-03	1.2E-02	0.42-1.24
Medical history										
Zhang 2017	Diabetes Mellitus	DM vs no DM	17	5036/ 2868215	1.22 (1.16-1.27)	1.32 (1.14-1.52)	1.23 (1.15-1.32)	1.4E-16	1.7E-04	0.81-2.15
Bernatsky 2011	SLE	Observed vs expected	4	44/40855	0.73 (0.53-1.00)	0.73 (0.53-1.00)	0.82 (0.54-1.20)	4.9E-01	4.9E-02	0.36-1.46
Li 2019	Endometriosis	Any vs none	25	1921/475988	1.58 (1.44-1.73)	2.35 (1.77-3.12)	1.34 (1.16-1.55)	8.82E-24	3.10E-09	0.87-6.34
Use of medical/ hormonal therapy										
Zhou 2008	HRT	Current vs ever	5	3958/ 1342899	1.25 (1.15-1.36)	1.28 (1.15-1.42)	1.20 (1.09-1.32)	8.5E-08	6.5E-06	1.01-1.62
Shi 2015	HRT	Ever vs never (cont E+P)	4	3337/ 1265735	1.22 (1.06-1.40)	1.22 (1.06-1.40)	1.13 (0.96-1.34)	5.8E-03	5.8E-04	0.90-1.65
Shi 2015	HRT	Ever vs never (seq E+P)	4	3337/ 1265735	1.23 (1.08-1.40)	1.35 (1.06-1.72)	1.14 (0.98-1.32)	1.4E-03	1.5E-02	0.54-3.35

Shi 2015	HRT	Ever vs never ET only	9	7512/ 2302683	1.40 (1.28-1.54)	1.44 (1.25-1.66)	1.31 (1.11-1.54)	2.8E-13	7.1E-07	0.99-2.09
Shi 2015	HRT	Ever vs never ET+PT	9	7512/ 2302683	1.28 (1.19-1.38)	1.23 (1.08-1.14)	1.50 (1.34-1.68)	1.1E-11	2.3E-03	0.87-1.75
Shi 2015	HRT	Ever vs never ET +E/PT	2	543/ 141880	1.55 (0.05-2.30)	1.55 (1.05-2.30)	1.50 (0.92-2.44)	2.7E-02	2.7E-02	NA
Coll Group 2015	HRT; Prospective	Current/recent vs never	11	11664/948390	1.37 (1.27-1.48)	1.37 (1.27-1.48)	1.28 (1.14-1.44)	1.3E-15	1.3E-15	1.26-1.50
Coll Group 2015	HRT; Prospective	Ever vs never	17	12110/950663	1.20 (1.13-1.28)	1.20 (1.13-1.28)	1.15 (1.06-1.26)	2.1E-09	2.1E-09	1.13-1.28
Coll Group 2015	HRT; Prospective	Ever vs never (info duration of use and times since last use)	14	11866/949657	1.24 (1.16-1.32)	1.24 (1.16-1.32)	1.16 (1.05-1.28)	5.2E-10	6.0E-10	1.15-1.33
Coll Group 2015	HRT; Prospective	Ever vs never	17	12110/950663	1.20 (1.13-1.28)	1.20 (1.13-1.28)	1.15 (1.06-1.26)	2.1E-09	2.1E-09	1.13-1.28
Coll Group 2008	OCP *	Ever vs never	45	7726/32201	0.74 (0.69-0.80)	0.74 (0.69-0.80)	0.74 (0.67-0.82)	5.7E-19	5.8E-16	0.68-0.81
Hankinson 1992	OCP	Ever vs never	3	60/ 80670	0.43 (0.25-0.75)	0.43 (0.25-0.75)	0.60 (0.30-1.40)	2.8E-03	3.0E-03	0.01-15.16
Baandrup 2013	NSAIDS; Non aspirin	Ever vs never	6	1782/ 505136	0.90 (0.81-1.00)	0.90 (0.81-1.00)	0.90 (0.75-1.08)	4.4E-02	4.4E-02	0.78-1.04
Wen 2019	Metformin	Ever vs never	3	3288/513702	0.16 (0.15-0.18)	0.18 (0.12-0.25)	0.16 (0.14-0.17)	0	2.5E-23	0.01-4.31
Reproductive factors										
Luan 2013	Breastfeeding	Per 5mo increase in duration	3	1180/ 447386	0.94 (0.90-0.99)	0.94 (0.89-1.00)	0.98 (0.92-1.05)	1.8E-02	3.4E-02	0.80-1.49
Siristatidis 2013	IVF	Ever vs never (reference group general population; excl OC diag <1yr post rx)	6	31606/ 1438001	1.50 (1.17-1.92)	1.47 (1.06-2.03)	1.30 (0.90-1.88)	1.2E-03	2.00E-02	0.73-2.96
Siristatidis 2013	IVF	Ever vs never (reference group IVF population; total follow up)	6	31606/ 1438001	1.62 (1.27-2.07)	1.66 (1.08-2.55)	1.35 (0.93-1.96)	1.0E-04	2.20E-02	0.52-5.28
Zhou 2017	PID	Ever vs never	6	8285/ 2929284	1.22 (1.11-1.34)	1.32 (1.05-1.66)	1.05 (0.92-1.20)	3.5E-05	1.62E-02	0.71-2.47
Environmental factors										
Reid 2011	Asbestos	Any vs none	14	5165/ 906145	1.77 (1.47-2.14)	1.86 (1.46-2.36)	1.30 (0.90-1.80)	2.4E-09	5.0E-07	1.05-3.29
Camargo 2011	Asbestos; Occupational	Total exposed vs nonexposed	20	126/ 21973	1.71 (1.42-2.06)	1.77 (1.37-2.27)	1.12 (0.66-1.80)	2.3E-08	9.7E-06	0.85-3.66
Camargo 2011	Asbestos; Occupational	High exposed vs nonexposed	6	20/ 6149	2.77 (1.69-4.54)	2.78 (1.36-5.66)	1.10 (0.37-2.21)	5.1E-05	5E-03	0.42-18.44

Abbreviations: BMI, body mass index; BMI iya, Body mass index in young adulthood; CC, case control; CI, confidence interval; CRP, C-reactive protein; cont E+P, continuous estrogen and progesterone; seq E+P, sequential estrogen and progesterone; ET, estrogen therapy; PT, progesterone therapy; E/PT, estrogen/ progesterone therapy; HvL, highest versus lowest; HR, hazard ratio; HRT, hormone replacement therapy; IVF, in vitro fertilisation; NA: Not available, due to <3 included studies; NSAID, non-steroidal anti-inflammatory drug; OC, ovarian cancer; OCP, oral contraceptive pill; PID, pelvic inflammatory disease; PoMP, postmenopausal; PrMP, premenopausal; RR, relative risk; rx, treatment; SLE, systemic lupus erythematous; T1T2, Type 1 or Type 2 diabetes mellitus; WG, weight gain

Key:

[α] Number of studies

[β] Fixed effects refers to summary relative risk 95% CI) using the meta-analysis fixed- effects model

[χ] Random effects refers to summary relative risk (95% CI) using the meta-analysis random -effects model.

[δ] Relative risk and 95% confidence interval of largest study (smallest SE) in each meta- analysis

[ε] P value of summary fixed effects estimate

[ϕ] P value of summary random effects estimate

[γ] Prediction intervals are reported only for meta-analyses including at least 3 studies

[η] Person years

All statistical tests were two-sided

***** %reduction in the standard error

Table 3.3: Evaluation of heterogeneity, small study effects and excess significance bias in the 40 meta-analyses investigating the risk factors associated with ovarian cancer incidence or mortality[§] –cohort studies.

Author, year	Exposure	Exposure contrast	Egger's P[α]	I^2 (95% CI) P[β]	Studies	Observed[χ]	Expected[δ], P-value[ε]					
							Fixed effects		Random effects		Largest study	
Anthrometric indices												
Pooralajal 2014	BMI	BMI ≥30kg/m^2 vs normal, PrMP	0.61	0 (0 -73) 0.86	3	1	0.68	0.66	0.68	0.66	0.66	0.64
Pooralajal 2014	BMI	BMI ≥30kg/m^2 vs normal	0.88	12 (0-54) 0.33	13	3	5.20	NP	5.20	NP	5.30	NP
Aune 2015	BMI	per 5kg/m2 increase	0.07	48 (6-67) 0.00	12	5	2.93	0.55	3.86	0.23	1.80	<0.01
Aune 2015	BMI	iya per 5kg/m2 increase	0.60	0 (0-61) 0.56	24	6	2.00	NP	2.00	NP	2.48	NP
Aune 2015	Weight	Per 5kg	0.42	7 (0-70) 0.36	6	1	0.23	0.10	0.24	0.10	0.21	0.08
Aune 2015	Height	Per 10cm	0.18	27 (0-59) 0.15	4	1	6.51	NP	6.86	NP	6.22	NP
Keum 2015	Weight gain	Per 5kg; never HRT (PoMP)	NA	NA	16	9	0.21	0.69	0.21	0.69	0.27	0.13
Dietary intake												
Song 2017	Calcium intake	HvL	0.62	0 (0-64) 0.61	5	0	1.73	NP	1.73	NP	1.66	NP
Larsson 2006	Dairy	HvL, total dairy products	NA	NA	2	1	1.91	NP	1.91	NP	1.87	NP
Larsson 2006	Dairy	HvL, skim/low fat	0.21	0 (0-73) 0.64	3	0	2.14	NP	2.14	NP	1.93	NP
Larsson 2006	Dairy	HvL, lactose	0.42	0 (0-73) 0.92	3	1	2.61	NP	2.61	NP	2.64	NP
Kolahdooz 2010	Meat	HvL, processed	0.37	0 (0-73) 0.93	3	0	1.79	NP	1.79	NP	1.56	NP
Wallin 2011	Meat	Per 100g/wk increment; red and processed	0.11	0 (0-41) 1.00	21	0	1.13	NP	1.13	NP	1.14	NP
WCRF CUP 2013	Non starchy vegetables	Per 100g/day	0.21	28 (0-71) 0.23	6	1	0.49	0.44	0.58	0.57	0.30	0.19
Butler 2011	Tea; black	HvL	0.44	15 (0-69) 0.32	5	2	3.88	NP	3.84	NP	4.71	NP
Zhang 2018	Non herbal tea	HvL	0.03	0 (0-73) 0.53	3	1	2.44	NP	2.44	NP	2.74	<0.01
Medical History												
Zhang 2017	Diabetes Mellitus	DM vs no DM	0.30	80 (67-86) 0.00	17	6	6.33	NP	8.29	NP	6.71	NP
Bernatsky 2011	SLE	Observed vs expected	0.97	0 (0-68) 0.53	4	0	0.36	NP	0.36	NP	0.26	NP
Li 2019	Endometriosis	Any vs none	0.00	83 (73-89) 0.00	25	12	9.49	0.18	13.70	NP	5.70	<0.01
Use of medical/ hormonal therapy												
Zhou 2008	HRT	Current vs ever	0.08	14 (0-69) 0.33	5	3	3.39	NP	3.61	NP	2.91	0.93
Shi 2015	HRT	Ever vs never (cont E+P)	0.07	0 (0-68) 0.50	4	1	2.55	NP	2.55	NP	1.63	NP
Shi 2015	HRT	Ever vs never (seq E+P)	0.18	50 (0-82) 0.11	4	2	2.70	NP	3.53	NP	1.75	0.80

Shi 2015	HRT	Ever vs never ET only	0.71	48 (0-74) 0.54	9	6	8.04	NP	8.28	NP	7.01	NP
Shi 2015	HRT	Ever vs never ET+PT	0.40	53 (0-76) 0.03	9	3	6.55	NP	5.60	NP	8.56	NP
Shi 2015	HRT	Ever vs never ET +E/PT	NA	NA	2	0	1.94	NP	1.94	NP	1.89	NP
Collab Gr 2015	HRT; Prospective	Current/recent vs never	0.68	0 (0-50) 0.78	11	3	8.72	NP	8.72	NP	7.20	NP
Collab Gr 2015	HRT- Prospective	Ever vs never	0.71	0 (0-45) 0.81	17	4	6.27	NP	6.27	NP	4.90	NP
Collab Gr 2015	HRT- Prospective	Ever vs never (info duration of use and times since last use)	0.97	0 (0-48) 0.45	14	2	6.71	NP	6.72	NP	4.87	NP
Wen 2019	Metformin	Ever vs never	0.38	14 (0-77) 0.31	3	2	2.41	NP	2.38	NP	2.42	NP
Zhou 2008	HRT	Current vs ever	0.08	14 (0-69) 0.32	5	3	3.39	NP	3.61	NP	2.91	0.93
Hankinson 1992	OCP	Ever vs never	0.41	NA	22	1	1.58	NP	1.58	NP	0.73	0.71
Baandrup 2013	NSAIDS Non aspirin	Ever vs never	0.55	0 (061) 0.93	6	0	1.02	NP	1.02	NP	1.02	NP
Collab Gr 2008	OCP	Ever vs never	0.61	0 (0-56) 0.61	45	3	6.67	NP	6.67	NP	6.68	NP
Reproductive factors												
Luan 2013	Breastfeeding	Per 5mo increase in duration	0.59	22 (0-78) 0.28	3	1	0.28	0.16	0.30	0.18	0.17	0.04
Siristatidis 2013	IVF	Ever vs never (reference group general population; excl OC diag <1yr post rx)	0.64	65 (0-83) 0.01	6	2	1.96	NP	1.89	NP	1.55	NP
Siristatidis 2013	IVF	Ever vs never (reference group IVF population; total follow up)	0.91	23 (0-69) 0.26	6	1	2.22	NP	2.30	NP	1.64	0.74
Zhou 2017	Pelvic inflammatory disease	Ever vs never	0.35	NA	3	1	9.88	0.25	13.7	NP	6.42	<0.01
Environmental factors												
Reid 2011	Asbestos[ζ]	Any vs none	0.64	28 (0 -61) 0.15	14	4	2.75	0.40	2.94	0.49	1.87	0.10
Camargo 2011	Asbestos; Occupational[ζ]	Total exposed vs nonexposed	0.72	35 (0-61) 0.06	20	6	4.11	0.59	5.14	NP	2.67	<0.01
Camargo 2011	Asbestos; Occupational[ζ]	High exposed vs nonexposed	0.78	45 (0-77) 0.12	6	2	2.93	0.55	3.86	0.23	1.80	<0.01

Abbreviations: BMI, body mass index; BMI iya, Body mass index in young adulthood; CC, case control; CI, confidence interval; CRP, C-reactive protein; cont E+P, continuous estrogen and progesterone; seq E+P, sequential estrogen and progesterone; ET, estrogen therapy; PT, progesterone therapy; E/PT, estrogen/ progesterone therapy; HvL, highest versus lowest; HR, hazard ratio; HRT, hormone replacement therapy; IVF, in vitro fertilisation; NA: Not available, due to <3 included studies; NP, not pertinent (P0.05<); NSAID, non-steroidal anti-inflammatory drug; OC, ovarian cancer; PoMP, postmenopausal; PrMP, premenopausal; RR, relative risk; rx, treatment; SLE, systemic lupus erythematous; T1T2, Type 1 or Type 2 diabetes mellitus; WG, weight gain

Key:

[α] P-value from the Egger's regression asymmetry test (P<0.10)

[β] I^2 metric of inconsistency (95% confidence interval) and the P-value of the Q test

[χ] Observed number of statistically significant studies in each meta-analysis

[δ] Expected number of statistically significant studies using the point estimate of each meta-analysis (from fixed effect, random effect of largest study accordingly) as the plausible effect size

[ε] P value of the excess statistical significance test

[ζ] Risk of dying from ovarian cancer

All statistical tests were two-sided

Table 3.4: A Measurement Tool to Assess Systematic Reviews 2 (AMSTAR 2) Summary quality assessment for all included systematic reviews

AMSTAR 2 Questions / Study Author, year	PICO	'A priori ' design and deviations justified	Study design	Literature search	Duplicate study selection review	Duplicate data extraction	Excluded studies	Description of included studies	Assess risk of bias	Funding	Statistical methods for meta-analysis	Impact of RoB from meta-analysis	RoB in individual studies in results	Heterogeneity	Small study bias	Conflict of interest	Score
Aune 2015	●	●	●	●	●	●	○	●	●	●	●	●	●	●	●	○	Low
Baandrup 2013	●	◎	●	○	○	●	○	●	●	●	●	○	●	●	●	●	Critically low
Barry 2014	●	●	●	●	●	●	○	●	●	●	●	○	●	●	●	●	Low
Berge 2017	●	●	●	●	●	●	○	●	●	○	●	●	●	●	●	●	Low
Bonovas 2006	●	◎	●	●	○	●	○	●	●	●	●	●	●	●	●	●	Low
Butler 2011	●	○	●	○	○	○	○	●	●	○	●	○	○	○	○	●	Critically low
Camargo 2011	●	◎	●	●	●	●	○	●	○	●	●	○	●	●	●	●	Critically low
Collabor group 2008	●	◎	●	●	○	○	○	●	◎	●	●	○	○	●	○	○	Critically low
Collabor group 2015	●	◎	●	●	●	●	○	●	●	●	●	●	●	●	○	○	Critically low
Gao 2013	●	○	●	◎	○	○	○	●	●	●	●	○	●	●	●	●	Critically low
Garg 1998	●	○	●	◎	○	○	○	●	○	○	●	○	○	●	●	○	Critically low
Gong 2013	●	●	●	◎	●	●	○	●	●	●	●	●	●	●	●	○	Low
Han 2014	●	◎	●	●	○	○	○	●	◎	○	●	●	●	●	●	○	Low
Huncharek 2001	●	●	○	●	○	●	◎	◎	●	○	●	●	●	●	●	●	Moderate
Huncharek 2007	●	○	○	◎	○	○	○	●	◎	●	●	○	●	●	○	○	Critically low
Huo 2018	●	◎	●	●	●	●	◎	●	●	○	●	●	●	●	●	○	Moderate
Ip 2007	●	◎	●	◎	○	○	○	●	●	●	●	○	●	●	●	○	Low
Jiang 2014	●	◎	●	●	●	●	◎	●	◎	●	●	●	●	●	●	●	High
Keum 2015 a	●	◎	●	◎	●	●	◎	●	●	○	●	○	●	●	●	○	Moderate
Keum 2015 b	●	●	●	◎	●	●	○	◎	●	○	●	●	●	●	●	●	Moderate
Kim 2010	●	○	●	◎	●	●	○	●	◎	○	●	○	○	●	●	●	Critically low
Kolahdooz 2010	●	◎	●	◎	○	○	○	●	●	●	●	●	●	●	●	●	Low
Koushik 2006	●	○	●	◎	○	○	○	◎	○	○	●	○	●	○	○	○	Critically low

Larsson 2006	●	○	●	◎	○	○	○	●	◎	●	●	○	●	●	●	○	Critically low
Lee 2016	●	◎	●	○	○	○	○	○	○	○	●	○	○	○	○	○	Critically low
Li 2019	●	◎	●	◎	●	●	◎	●	◎	○	●	○	●	●	●	●	High
Li 2017	●	◎	●	◎	●	●	◎	●	○	●	●	○	●	●	●	●	Low
Li 2014 a	●	◎	●	◎	○	●	◎	●	◎	●	●	●	●	●	●	○	Moderate
Li 2014 b	●	○	○	◎	○	●	●	●	○	○	○	●	●	●	●	●	Critically low
Liu 2019	●	◎	●	●	●	●	◎	●	●	●	●	●	●	●	●	●	High
Liu 2014	●	◎	●	●	○	●	○	●	●	○	●	●	●	●	●	●	Low
Luan 2013	●	○	●	◎	○	●	○	●	○	●	●	○	●	●	●	●	Critically low
Murphy 2012	●	○	○	○	○	○	○	●	○	○	●	○	●	●	○	○	Critically low
Negri 1991	○	○	○	○	○	○	○	○	○	○	○	○	○	●	○	○	Critically low
Olsen 2007	●	○	●	○	○	○	○	●	●	○	●	○	●	●	●	●	Critically low
Pang 2018	●	◎	●	◎	●	●	◎	●	●	●	●	●	●	●	●	●	High
Park 2016	○	○	○	○	○	○	○	○	○	○	●	○	○	○	○	○	Critically low
Pelucchi 2015	●	○	○	◎	○	○	○	●	○	●	●	○	○	○	○	●	Critically low
Penninkilampi 2018	●	●	●	◎	○	○	◎	●	●	○	●	○	●	●	●	●	Moderate
Poorolajal 2014	●	◎	●	◎	●	●	◎	●	●	●	●	●	●	●	●	●	High
Qiu 2016	●	●	●	○	○	○	●	●	●	●	●	○	●	●	●	●	Moderate
Qu 2014	●	◎	●	◎	●	●	○	●	●	●	●	●	●	●	●	●	Low
Reid 2011	●	○	○	○	○	○	○	●	●	○	●	○	○	○	○	●	Critically low
Salar 2019	●	◎	●	◎	○	●	◎	●	●	○	●	○	○	○	●	○	Low
Santucci 2019	●	●	●	●	○	○	○	●	○	○	●	○	○	●	●	○	Critically low
Schmid 2014	●	◎	●	◎	○	●	◎	●	◎	○	●	○	●	●	●	●	Moderate
Shafei 2019	●	◎	●	◎	●	●	◎	○	○	●	●	○	●	●	●	●	Low
Shi 2015	●	◎	○	◎	○	●	◎	●	◎	●	●	○	●	●	●	●	Moderate
Siristatidis 2013	●	●	●	◎	●	●	○	●	◎	●	●	●	●	●	○	●	Critically low
Song 2017	●	◎	●	◎	○	●	○	●	◎	○	●	○	○	●	○	●	Critically low
Song 2018	●	●	○	◎	●	●	○	◎	○	○	●	○	○	●	●	●	Critically low

Wallin 2011	●	○	●	◎	○	●	○	●	○	○	●	○	○	●	●	○	Critically low
Wang 2018	●	◎	●	◎	●	●	◎	●	◎	○	●	○	●	●	●	●	Moderate
Wen 2019	●	◎	●	◎	○	●	◎	●	◎	●	●	●	●	●	●	●	Moderate
Whiteman 2000*	●	○	○	○	○	○	●	●	◎	○	●	●	●	●	○	○	Low
Wu 2019	●	◎	●	◎	●	●	◎	●	◎	●	●	○	●	●	●	●	Moderate
Xu 2018	●	◎	●	◎	●	●	◎	●	◎	●	●	○	●	●	●	●	Moderate
Yan 2015	●	◎	●	◎	●	●	○	●	◎	○	●	○	●	●	●	●	Low
Yang 2019	●	●	○	●	○	○	○	◎	●	○	●	●	●	○	●	●	Moderate
Yin 2011	●	◎	●	◎	●	●	○	●	○	●	●	○	●	●	●	●	Critically low
Zeng 2015	●	◎	●	◎	●	●	◎	●	●	○	●	●	●	●	●	●	Moderate
Zeng 2016	●	◎	●	◎	●	●	◎	●	●	●	●	●	●	●	●	●	Moderate
Zhan 2017	●	◎	●	◎	○	○	◎	●	●	●	●	○	●	●	●	●	Moderate
Zhang 2016	●	◎	●	◎	●	●	◎	●	●	●	●	○	●	●	●	●	Moderate
Zhang 2017	●	◎	●	◎	●	●	◎	●	●	●	●	○	●	●	●	●	Moderate
Zhang 2018 a	●	◎	●	◎	●	●	◎	●	●	●	●	○	●	●	●	●	Moderate
Zhang 2018 b	●	◎	●	◎	●	●	○	●	●	○	●	○	●	●	●	●	Low
Zheng 2018	●	●	●	◎	○	●	○	●	●	●	●	○	●	●	●	○	Low
Zhong 2014	●	◎	●	◎	○	●	○	●	●	○	●	○	●	○	●	●	Low
Zhou 2014	●	◎	●	◎	●	●	○	●	●	○	●	○	●	○	●	○	Critically low
Zhou 2017	●	◎	●	◎	○	●	○	●	●	●	●	○	●	●	●	●	Low
Zhou 2008	●	◎	●	○	○	●	○	●	●	●	●	○	●	●	●	○	Critically low

Key: ● Yes

◎ Partial yes

○ No

* Pooled analysis

NA Not Applicable

Critical flaw

3.4.7 Grading the evidence

Each of the meta-analyses was graded as presenting strong, highly suggestive, suggestive or weak evidence (Tables 3.5 and 3.6)[157,159,232]. Six out of 174 meta-analyses (3%) satisfied the criteria of strong evidence for an association with ovarian cancer incidence; these included prospective studies. Height (RR per 10cm 1.16; 95%CI 1.11-1.20), BMI (RR ≥30kg/m^2 vs <25 kg/m^2 1.27; 95%CI 1.17-1.38), and three associations for use of HRT (ever versus never use - RR 1.20; 95%CI 1.13-1.28; current or recent use vs. never use – RR 1.37, 95%CI 1.27-1.48; ever vs. never use, including information on duration of use – RR 1.25, 95% CI 1.16-1.32) were associated with a higher risk of ovarian cancer. Use of OCP was strongly associated with lower risk of developing ovarian cancer (ever versus never, RR 0.74, 95%CI 0.69-0.80).

Only three meta-analyses (3/174, 2%) fulfilled the criteria for highly suggestive evidence. Two of them (2/174, 1%) reported an increased ovarian cancer incidence for use of HRT (ever versus never use, oestrogen therapy only, RR 1.44; 95% CI 1.25-1.66) and endometriosis (any versus none, RR 2.34; 95%CI 1.76-3.10). Metformin use was associated with a lower risk of ovarian cancer incidence (ever versus never, RR 0.18; 95%CI 0.12-0.25). There were no exposures that showed a strong or highly suggestive association with an increased risk of ovarian cancer mortality.

Five meta-analyses (5/174, 3%) received a suggestive evidence grade and 26 (26/174, 15%) meta-analyses had weak evidence, whereas the rest (134/174, 77%) had a nominal insignificant association.

Table 3.5: Summary of evidence grading for meta-analysis of risk factors associated with ovarian cancer incidence or mortality– cohort studies only*

Evidence	Criteria used	Decreased Risk	Increased risk
Strong	$P<10^{-6}$‖; >1,000 cases; $I^2<50\%$; no small study effects¶; prediction interval excludes the null value; no excess significance bias† **n=6**	***Use of medical/hormonal therapy*** OCP inc Ever vs never **	***Anthropometric measure*** Height inc Per 10cm BMI; ≥30kg/m^2 vs normal
			Use of medical/hormonal therapy HRT inc Ever vs never (prospective studies) HRT inc Current/recent vs never (prospective studies) HRT inc Ever vs never (prospective studies; info duration of use and time since last use known)
Highly Suggestive	$P<10^{-6}$‖; >1,000 cases; P<0.05 of the largest study in a meta-analysis n=3	***Use of medical/hormonal therapy*** Metformin inc Ever vs never	***Use of medical/hormonal therapy*** HRT inc Ever vs never ET only
			Medical history Endometriosis inc Any vs none
Suggestive	$P<10^{-3}$‖; >1000 cases n=5	***None***	***Anthropometric measures*** BMI inc iya per 5kg/m^2 increase BMI inc per 5kg/m^2 increase
			Asbestos Any vs none mor
			Medical history Diabetes inc; Yes vs no
			Use of medical/hormonal therapy HRT inc; Current vs ever
Weak	P<0.05 n=26	**Reproductive factors** Breastfeeding inc; Per 5mo increase in duration	**Anthropometric measures** BMI inc PrMP Obese vs normal

		Use of medical/ hormonal therapy NSAIDS inc, Non aspirin; Ever vs never OCP inc; Ever vs never	BMI inc PoMP; Obese vs normal Per 5kg weight inc WG per 5kg increase PoMP, HRT, inc
		Medical history SLE inc; observed vs expected	**Asbestos** Total exposed vs non- exposed, mor High exposed vs non- exposed, mor
		Dietary Intake Tea (black) inc; Highest vs lowest Non herbal tea inc; Highest vs lowest Calcium inc; Highest vs lowest Non-starchy vegetables inc; Per 100g/day	**Dietary intake** Dairy total inc; Highest vs lowest Dairy skim/low fat inc; Highest vs lowest Dairy lactose inc; Highest vs lowest Meat (processed) inc; Highest vs lowest Meat (red and processed) mor; Per 100g/week increment
			Use of medical/hormonal therapy HRT inc; Ever vs never (continuous E+P) HRT inc; Ever vs never (sequential E+P) HRT inc; Ever vs never (E+P) HRT inc; Ever vs never (E+E/P)
			Reproductive Factors PID inc; Ever vs never IVF inc; Ever vs never (reference group general population) IVF inc; Ever vs never (reference group IVF population)

Abbreviations: BMI, body mass index; BMI iya, body mass index in young adulthood; BMI PrMP, body mass index premenopausal; CC, case control; HRT, hormone replacement therapy; inc, incidence; mor, mortality; NSAID, non- steroidal anti -inflammatory drugs; WG, weight gain; BMI PoMP, body mass index postmenopausal; CRP, c-reactive protein; OCP, oral contraceptive pill; SLE, systemic lupus erythematous; E+P, estrogen and progesterone; E+E/P, estrogen and estrogen/progesterone; PID, pelvic inflammatory disease; IVF, in-vitro fertilisation; RR, relative risk

Key:

*only meta-analyses meeting at least weak grade of evidence listed

**% reduction of the standard error

‖P indicates the p-values of the meta-analysis random effects model

ϥSmall study effect is based on the P-value from the Egger's regression asymmetry test (P>0.1) where the random effects summary estimate was larger compared to the point estimate of the largest study in a meta-analysis

†Based on the p-value (P>0.1) of the excess significance test using the largest study (smallest standard error) in a meta-analysis as the plausible effect size.

Table 3.6: Details of evidence grading for meta-analysis of risk factors for ovarian cancer incidence or mortality – only cohort studies included*

Exposure	Exposure contrast	N*	Sample size Cases/cohort	Largest study#	Random effects summary RR (95% CI)¥	Random P-value‖	95% Prediction interval	Egger's P∞	I^2 (%)	Excess significance$		Evidence Grading**
										O/E α	P-value Φ	
Strong evidence												
Anthropometric measure												
Height	Per 10cm	16	18663/13600000	1.14 (1.10-1.18)	1.16 (1.11-1.20)	2.2E-13	1.06-1.26	0.18	27	9/6.22	0.15	Strong
BMI	Obese vs normal	13	6947/20,560,388 ι	1.27 (1.19-1.36)	1.27 (1.17-1.38)	2.6E-8	1.09-1.47	0.88	12	3/5.30	NP	Strong
Use of medical/hormonal therapy												
HRT- Prospective	Current/recent vs never	12	11664/948390	1.28 (1.14-1.44)	1.37 (1.27-1.48)	1.3E-15	1.26-1.50	0.68	0	3/7.20	NP	Strong
HRT- Prospective	Ever vs never	17	12110/950663	1.15 (1.06-1.26)	1.20 (1.13-1.28)	2.1E-09	1.13-1.28	0.71	0	4/4.90	NP	Strong
HRT- Prospective	Ever vs never (info duration of use and time since last use)	14	11866/949657	1.16 (1.05-1.28)	1.24 (1.16-1.32)	6.0E-10	1.15-1.33	0.97	0	2/4.87	NP	Strong
OCP	Ever vs never	45	7726/32201	0.74 (0.67-0.82)	0.74 (0.69-0.80)	5.8E-16	0.68-0.81	0.61	0	3/6.68	NP	Strong
Highly suggestive evidence												
Use of medical/hormonal therapy												
HRT	Ever vs never ET only	11	7512/2302683	1.31 (1.11-1.54)	1.44 (1.25-1.66)	7.1E-07	0.99-2.09	0.71	48	6/7.01	NP	Highly suggestive
Metformin	Ever vs never	3	3288/513702	0.16 (0.14-0.17)	0.18 (0.12-0.25)	2.5E-23	0.01-4.31	0.38	14	2/2.42	NP	Highly suggestive
Anthropometric measure												
Height	per 5cm increase	13	16198/3514114	1.07 (1.05-1.09)	1.07 (1.05-1.10)	1.2E-09	1.02-1.14	0.31	32	8/2.77	<0.01	Highly suggestive
Medical history												
Metformin	Ever vs never	3	3288/513702	0.16 (0.14-0.17)	0.18 (0.12-0.25)	2.5E-23	0.01-4.31	0.38	14	2/2.42	NP	Highly suggestive
Suggestive evidence												
Environmental factors												
Asbestos	Any vs none	14	5165/906145	1.30 (0.90-1.80)	1.86 (1.46-2.36)	5.0E-07	1.05-3.29	0.64	28	4/1.87	0.10	Suggestive
Anthropometric measures												
BMI	per 5kg/m^2 increase	24	17734/16300000	0.97 (0.93-1.01)	1.07 (1.04-1.11)	<0.01	0.96-1.21	0.07	48	6/1.80	<0.01	Suggestive
BMI	iya per 5kg/m^2 increase	6	9452/11100000	1.16 (1.04-1.29)	1.12 (1.05-1.19)	<0.01	1.03-1.23	0.60	0	0/1.00	NP	Suggestive
Medical history												
Diabetes Mellitus	DM vs no DM	17	5036/2868215	1.23 (1.15-1.32)	1.32 (1.14-1.52)	<0.01	0.81-2.15	0.30	80	6/6.71	NP	Suggestive
Use of medical / hormonal therapy												

HRT	Current vs ever	5	3958/1342899	1.20 (1.09-1.32)	1.28 (1.15-1.42)	6.5E-06	1.01-1.62	0.08	14	3/2.91	0.93	Suggestive
Weak evidence												
Asbestos												
Asbestos	Total exp vs nonexp	20	126/21973	1.12 (0.66-1.80)	1.77 (1.37-2.27)	9.7E-06	0.85-3.66	0.72	35	6/1.06	<0.01	Weak
Asbestos	High exp vs nonexp	6	20/6149	1.10 (0.37-2.21)	2.78 (1.36-5.66)	0.01	0.42-18.44	0.78	45	2/0.31	<0.01	Weak
Anthropometric measure												
BMI	PrMP Obese vs normal	3	71/350211	1.56 (1.14-2.16)	1.57 (1.20-2.06)	<0.01	0.27-9.02	0.61	0	1/0.66	0.64	Weak
BMI	PoMP Obese vs normal	5	350/546195	1.02 (0.82-1.26)	1.23 (1.03-1.47)	2.6E-01	0.72-2.09	0.52	46	1/ 0.25	0.13	Weak
Weight	Per 5kg weight	4	1006/297350	1.02 (1.00-1.05)	1.03 (1.01-1.05)	<0.01	0.98-1.08	0.42	7	1/0.21	0.08	Weak
Weight gain	per 5kg increase PoMP, HRT	2	217/23984	1.16 (1.03-1.31)	1.13 (1.03-1.24)	0.01	NA	NA	NA	1/0.27	0.13	Weak
Dietary intake												
Dairy, total products	Highest vs lowest	2	427/90001	1.61 (1.07-2.42)	1.66 (1.19-2.31)	<0.01	NA	NA	NA	1/1.87	NP	Weak
Dairy, skim/low fat	Highest vs lowest	3	728/170327	1.32 (0.97-1.82)	1.35 (1.09-1.68)	<0.01	0.35-5.43	0.21	0	0/1.93	NP	Weak
Dairy, lactose	Highest vs lowest	3	728/170327	1.48 (1.05-2.09)	1.47 (1.17-1.84)	<0.01	0.34-6.29	0.42	0	1/ 2.64	NP	Weak
Meat; processed	Highest vs lowest	3	1018/696100	1.23 (0.92-1.63)	1.26 (1.02-1.56)	3.5E-02	0.31-5.07	0.37	0	0/ 1.56	NP	Weak
Meat; red and processed	Per 100g/week increment	21	6536/2140286	1.02 (0.98-1.06)	1.01 (1.00-1.04)	3.4E-02	1.00-1.04	0.12	0	0/ 1.14	NP	Weak
Non starchy vegetables	Per 100g/day	6	2053/641079	1.00 (0.93-1.07)	0.94 (0.89-1.00)	4.0E-02	0.82-1.08	0.21	28	1/0.30	0.19	Weak
Tea; black	Highest vs lowest	5	1299/203998	0.63 (0.40-0.99)	0.73 (0.56-0.93)	1.2E-02	0.42-1.24	0.44	15	2/4.71	NP	Weak
Calcium	Highest vs lowest	5	1726/351192	0.86 (0.68-1.10)	0.86 (0.74-1.00)	4.0E-02	0.67-1.09	0.62	0	0/1.66	NP	Weak
Non herbal tea	Highest vs lowest	3	734/164882	0.63 (0.40-0.99)	0.69 (0.52-0.93)	1.4E-02	0.11-4.57	0.03	0	1/ 2.74	NP	Weak
Use of medical / hormonal therapy												
HRT	Ever vs never (continuous E+P)	4	3337/1265735	1.13 (0.96-1.34)	1.22 (1.06-1.40)	<0.01	0.90-1.65	0.07	0	1/1.63	NP	Weak
HRT	Ever vs never (sequential E+P)	4	3337/1265736	1.14 (0.98-1.32)	1.35 (1.06-1.72)	1.5E-02	0.54-3.35	0.18	50	2/1.75	0.80	Weak
HRT	Ever vs never (ET+PT)	9	7512/2302683	1.50 (1.34-1.68)	1.23 (1.08-1.14)	2.3E-03	0.87-1.75	0.40	53	3/8.65	NP	Weak
HRT	Ever vs never (ET+ E/PT)	2	543/141880	1.50 (0.92-2.44)	1.55 (1.05-2.30)	2.7E-02	NA	NA	NA	0/1.89	NP	Weak
NSAIDS Non aspirin	Ever vs never	6	1782/505136	0.90 (0.75-1.08)	0.90 (0.81-1.00)	4.4E-02	0.78-1.04	0.55	0	0/1.02	NP	Weak
OCP	Ever vs never	3	60/80670	0.60 (0.30-1.40)	0.43 (0.25-0.75)	3.0E-03	0.01-15.16	0.41	NA	1/0.73	0.71	Weak
Reproductive factors												

PID	Ever vs never	6	8285/2929284	1.05 (0.92-1.20)	1.32 (1.05-1.66)	1.6E-02	0.71-2.47	0.35	65	2/1.01	0.28	Weak
IVF	Ever vs never (reference group general population excluding OC diagnosis <1yr post treatment)	6	31606/1438001	1.30 (0.90-1.88)	1.47 (1.06-2.03)	2.0E-02	0.73-2.96	0.64	23	1/1.55	NP	Weak
IVF	Ever vs never reference group IVF population; total follow up	6	31606/1438002	1.35 (0.93-1.96)	1.66 (1.08-2.55)	2.2E-02	0.52-5.28	0.91	52	2/1.64	0.74	Weak
Breastfeeding	Per 5mo increase in duration	3	1180/ 447386	0.98 (0.92-1.05)	0.94 (0.89-1.00)	0.03	0.80-1.49	0.59	22	1/0.17	0.04	Weak
Medical history												
SLE	Observed vs expected	4	44/40855	0.82 (0.54-1.20)	0.73 (0.53-1.00)	4.9E-02	0.36-1.46	0.97	0	0/0.26	NP	Weak

Abbreviations: BMI, body mass index; BMI iya, body mass index in young adulthood; BMI PoMP, body mass index postmenopausal; BMI PrMP, body mass index premenopausal; CC, case control; CRP, c-reactive protein; E+P, estrogen and progesterone; E+E/P, estrogen and estrogen/progesterone; HRT, hormone replacement therapy; IVF, in-vitro fertilisation; NA, not available; np; not pertinent, because the estimated is larger than the observed, and there is no evidence of excess statistical significance based on the assumption made for the plausible effect size; NSAID, non- steroidal anti -inflammatory drugs; OC, ovarian cancer; OCP, oral contraceptive pill; PID, pelvic inflammatory disease; RR, relative risk; SLE, systemic lupus erythematous; WG, weight gain

Key:

*****only meta-analyses meeting at least weak grade of evidence listed

* Number of studies

Relative risk and 95% confidence interval of largest study (smallest standard error) in each meta-analysis

¥ Random effects refer to summary risk ratio (95% confidence interval) using the random-effects model

|| P value of summary random effects estimate

" P-value from the Egger's regression asymmetry test

§ Expected number of statistically significant studies using the point estimate of the largest study (smallest standard error) as the plausible effect size

α Observed/Expected number of statistically significant studies

φP value of the excess statistical significance test

All statistical tests were two-sided

[q]Small study effect is based on the P-value from the Egger's regression asymmetry test (P>0.1) where the random effects summary estimate was larger compared to the point estimate of the largest study in a meta-analysis

†Based on the p-value (P>0.1) of the excess significance test using the largest study (smallest standard error) in a meta-analysis as the plausible effect size.

ι Person years

**Summary of evidence grading criteria:

Weak	$P<0.05$[ll]
Suggestive	$P<10^{-3}$[ll]; >1,000 cases
Highly suggestive	$P<10^{-6}$[ll]; >1,000 cases; P<0.05 of the largest study in a meta-analysis
Strong	$P<10^{-6}$[ll]; >1,000 cases; P<0.05 of the largest study in a meta-analysis; $I^2<50\%$; no small study effect[q]; prediction interval excludes the null value; no excess significance bias[†]

Table 3.7: Details of evidence grading for meta-analyses of risk factors for ovarian cancer incidence or mortality[κ] – all statistically significant study types included*

Exposure	Exposure contrast	N*	Sample size Cases/Cohort	Largest study#	Random effects summary RR (95% CI) ¥	Random P-value||	95% Prediction interval	Egger's P	I2 (%)	Excess significance		Evidence Grading**
											O/E ¤	P-value	
Strong evidence													
Anthropometric indices													
Height	per 10cm	16	18663/13600000	1.14 (1.10-1.18)	1.16 (1.11-1.20)	2.2E-13	1.06-1.26	0.18	27	9/6.22	0.15	1: Strong
BMI	BMI ≥30kg/m² vs normal	13	6947/ 20,560,388ι	1.27 (1.19-1.36)	1.37 (1.17-1.38)	2.6E-8	1.09-1.47	0.88	12	3/5.30	NP	1: Strong
Use of medical or hormonal therapy												
HRT- Prospective	Ever vs never	17	12110/950663	1.15 (1.06-1.26)	1.20 (1.13-1.28)	2.1E-09	1.13-1.28	0.71	0	4/4.00	NP	1: Strong
HRT - Prospective	Ever vs never (info duration of use and times since last use)	14	11866/949657	1.16 (1.05-1.28)	1.24 (1.16-1.32)	6.0E-10	1.15-1.33	0.97	0	2/4.87	NP	1: Strong
HRT- Prospective	Cur/rec vs never	12	11664/948390	1.28 (1.14-1.44)	1.37 (1.27-1.48)	1.3E-15	1.26-1.50	0.68	0	3/7.20	NP	1: Strong
Repro factors: OCP	Ever vs never	45	19610/90211	0.74 (0.67-0.82)	0.74 (0.70-0.78)	9.0E-30	0.70-0.78	0.16	0	11/23	NP	1: Strong
Highly suggestive evidence												
Reproductive factors												
Breastfeeding	Ever vs Never	55	17139/415985	0.53 (0.46-0.62)	0.69 (0.34-0.76)	9.6E-16	0.43-1.13	0.71	76	24/3.85	<1e-100	2: Highly suggestive
Use of medical or hormonal therapy												
HRT	Ever vs never; ET only	11	7512/2302683	1.31 (1.11-1.54)	1.44 (1.25-1.66)	7.1E-07	0.99-2.09	0.71	48	6/7.01	NP	2: Highly suggestive
Metformin	Ever vs never	3	3288/513702	0.16 (0.14-0.17)	0.18 (0.12-0.25)	2.5E-23	0.01-4.31	0.38	NA	2/2.42	NP	2: Highly suggestive
Medical history												
Endometriosis	Any vs none	25	15921/858822	1.46 (1.31-1.63)	1.99 (1.67-2.32)	1.3E-17	1.05-3.75	0.00	76	22/17.50	0.08	2: Highly suggestive
Carcinogens/ environmental irritants												
Genital talc powder use	Ever vs never	27	15154/367198	1.32 (1.14-1.50)	1.26 (1.17-1.35)	1.3E-10	1.00-1.59	0.70	39	13/17.20	NP	2: Highly suggestive
Suggestive evidence												
Reproductive factors												
Tubal ligation	Ever vs never	25	31678/373632	0.81 (0.73-0.91)	0.70 (0.60-0.81)	1.3E-06	0.35-1.38	0.78	86	16/14.30	0.5	3: Suggestive
Use of medical or hormonal therapy												
HRT	Current vs ever	5	3958/1342899	1.20 (1.09-1.32)	1.28 (1.15-1.42)	6.5E-06	1.01-1.62	0.08	14	3/2.91	NP	3: Suggestive
Anthropometric indices												
BMI	per 5kg/m2 increase	25	17734/16300000	0.97 (0.93 1.01)	1.07 (1.04-1.11)	<0.01	0.96-1.21	0.07	48	6/1.80	0.00	3: Suggestive
BMI	iya per 5kg/m2 increase	6	9452/11100000	1.16 (1.04 1.29)	1.12 (1.05 -1.19)	<0.01	1.03-1.23	0.60	0	1/2.48	NP	3: Suggestive
Carcinogen/ environmental irritants												
Asbestosᵏ	Any vs none	16	5165/906145	1.30 (0.90-1.80)	1.86 (1.46-2.36)	5.0E-07	1.05 -3.29	0.64	28	4/1.87	0.1	3: Suggestive
Perineal talc use	Long term (>10yrs) vs none	12	2125/134414	1.29 (1.05-1.59)	1.25 (1.10-1.43)	<0.01	0.88-1.79	0.30	45	4/1.05	0.53	3: Suggestive
Medical history												
Diabetes Mellitus	DM vs no DM	17	5036/2868215	1.23 (1.15-1.32)	1.32 (1.14 -1.52)	<0.01	0.81-2.15	0.30	80	6/6.71	NP	3: Suggestive
Dietary factors												
Polyphenol; Genistein	HvL	4	1239/99776	0.42 (0.30-0.60)	0.56 (0.42-0.76)	<0.01	0.18-1.72	0.26	46	2/4.00	NP	3: Suggestive

Polyphenol; Soyfoods	HvL	6	2617/180668	0.54 (0.43-0.69)	0.51 (0.39-0.68)	2.9E-06	0.22-1.23	0.84	70	4/5.82	NP	3: Suggestive
Polyphenol; Tofu	HvL	4	1111/162722	0.57 (0.40-0.80)	0.59 (0.47-0.75)	<0.01	0.35-1.00	0.87	0	2/3.76	NP	3: Suggestive
Dietary fibre intake	Per 10g/day increase	9	3492/137025	0.91 (0.81-1.01)	0.91 (0.81-1.01)	5.0E-05	0.79-0.97	0.58	7	3/1.31	0.11	3: Suggestive
Dietary inflammatory index	HvL	4	3104/7982	1.31 (1.06-1.62)	1.31 (1.06-1.62)	8.6E-06	1.01-1.98	0.39	0	3/3.07	NP	3: Suggestive
Biochemical markers												
CRP	>10mg/L vs <10mg/L	4	1413/31442	1.67 (1.03-2.70)	2.14 (1.47-3.11)	<0.01	0.77-5.93	0.15	13	2/3.85	NP	3: Suggestive
Weak evidence												
Reproductive factors												
Breastfeeding	Per 5mo increase in duration	25	1180/447386	0.98 (0.92-1.05)	0.94 (0.89-1.00)	0.03	0.80-1.49	0.59	22	1/0.17	0.04	4: Weak
Breastfeeding	>12mo vs never	6	2833/11659	1.08 (0.87-1.34)	0.72 (0.53-0.97)	0.03	0.29-1.82	0.30	71	2/0.74	0.12	4: Weak
PID	Ever vs never	14	8285/2929284	1.05 (0.92-1.20)	1.32 (1.05-1.66)	1.6E-02	0.71-2.47	0.35	65	2/1.01	0.28	4: Weak
Age at menarche	Oldest vs youngest	24	8872/1022451	0.77 (0.62-0.95)	0.85 (0.75-0.97)	0.01	0.57-1.28	0.87	42	4/1.31	<0.01	4: Weak
IVF	Ever vs never (ref group gen population excl OC diagnosis <1yr post treatment)	6	31606/1438001	1.30 (0.90-1.88)	1.47 (1.06-2.03)	2.0E-02	0.73-2.96	0.64	23	1/1.55	NP	4: Weak
IVF	Ever vs never ref group IVF pop; total follow up	6	31606/1438002	1.35 (0.93-1.96)	1.66 (1.08-2.55)	2.2E-02	0.52-5.28	0.91	52	2/1.64	0.74	4: Weak
Nulliparity	Yes vs no	3	285/835	1.40 (1.10-1.70)	1.42 (1.18-1.72)	<0.01	0.42-4.84	0.87	0	1/0.89	0.89	4: Weak
Age at last birth	Highest vs lowest	13	19959/2470521	0.93 (0.85-1.01)	0.77 (0.65-0.91)	2.2E-03	0.45-1.32	0.58	74	5/2.37	0.06	4: Weak
Use of medical or hormonal therapy												
OCP	Ever vs never	22	60/80670	0.60 (0.30-1.40)	0.43 (0.25-0.75)	3.0E-03	0.01-15.16	0.41	13	1/0.73	0.71	4: Weak
Aspirin	Ever vs never	22	14581/511293	0.94 (0.85-1.04)	0.89 (0.83-0.96)	<0.01	0.75-1.06	0.01	23	3/2.74	0.86	4: Weak
Aspirin	Ever vs never CC	14	12593/101479	0.94 (0.85-1.04)	0.85 (0.77-0.94)	<0.01	0.68-1.10	0.01	38	3/2.07	0.48	4: Weak
NSAIDS, Non aspirin	Ever vs never	16	1782/505136	0.90 (0.75-1.08)	0.90 (0.81-1.00)	4.4E-02	0.78-1.04	0.55	0	0/1.02	NP	4: Weak
Statin use	Any vs none long-term	3	231/1829	0.54 (0.27-1.09)	0.48 (0.28-0.80)	<0.01	0.02-13.95	0.31	0	1/2.04	NP	4: Weak
Statin use	Any vs none	5	648/84234	0.83 (0.66-1.05)	0.79 (0.64-0.98)	0.03	0.57-1.11	0.23	0	0/1.88	NP	4: Weak
Antidepressants	Ever vs never	19	8200/571701	0.99 (0.83-1.18)	0.70 (0.57-0.87)	<0.01	0.28-1.75	0.28	83	7/0.97	<1e-100	4: Weak
HRT	Invasive; ever vs never	10	4392/255821	1.15 (0.94-1.42)	1.16 (1.01-1.32)	0.04	0.84-1.59	0.17	32	2/2.82	NP	4: Weak
HRT	Invasive and borderline; ever vs never	12	5421/260387	1.15 (0.94-1.42)	1.13 (1.01-1.27)	0.03	0.90-1.43	0.14	21	2/3.36	NP	4: Weak
HRT	Ever vs never (cont E+P)	4	3337/1265735	1.13 (0.96-1.34)	1.22 (1.06-1.40)	<0.01	0.90-1.65	0.07	0	1/1.63	NP	4: Weak
HRT	Ever vs never (seq E+P)	4	3337/1265735	1.14 (0.98-1.32)	1.35 (1.06-1.72)	1.5E-02	0.54-3.35	0.18	51	2/1.75	0.80	4: Weak
HRT	Ever vs never ET+PT	11	7512/2302683	1.50 (1.34-1.68)	1.23 (1.08-1.14)	2.3E-03	0.87-1.75	0.34	53	3/8.56	NP	4: Weak
HRT	Ever vs never ET +E/PT	2	543/141880	1.50 (0.92-2.44)	1.55 (1.05-2.30)	2.7E-02	NA	NA	0	0/1.89	NP	4: Weak
Anthropometric indices												
BMI	PrMP BMI ≥30kg/m^2 vs normal	3	71/350211	1.56 (1.14-2.16)	1.57 (1.20-2.06)	<0.01	0.27-9.02	0.61	0	1/0.66	0.64	4: Weak
BMI	PoMP; BMI ≥30kg/m^2 vs normal	5	350/546195	1.02 (0.82-1.26)	1.23 (1.03-1.47)	2.6E-01	0.72-2.09	0.52	46	1/0.25	0.13	4: Weak

BMI	BMI Highest vs lowest	9	1801/335870	1.20 (0.90-1.50)	1.20 (1.00-1.44)	0.05	0.74-1.95	0.37	47	3/1.63	0.24	4: Weak
Weight	Per 5kg weight	4	1006/297350	1.02 (1.00-1.05)	1.03 (1.01-1.05)	<0.01	0.98-1.08	0.42	7	1/0.21	0.08	4: Weak
Weight gain	per 5kg increase PoMP, HRT	2	217/23984	1.16 (1.03-1.31)	1.13 (1.03-1.24)	<0.01	NA	NA	NA	1/0.27	0.13	4: Weak
Carcinogen/ environmental irritant												
Asbestos; Occupational[κ]	Total exposed vs nonexposed	20	126/21973	1.12 (0.66-1.80)	1.77 (1.37-2.27)	9.7E-06	0.85-3.66	0.72	35	6/1.06	<0.01	4: Weak
Asbestos; Occupational[κ]	High exposed vs nonexposed	6	20/6149	1.10 (0.37-2.21)	2.78 (1.36-5.66)	<0.4	0.42-18.44	0.78	45	2/0.31	<0.01	4: Weak
Pre existing medical conditions/ interventions												
SLE	Observed vs expected	5	44/40855	0.82 (0.54-1.20)	0.73 (0.53-1.00)	4.9E-02	0.36-1.46	0.97	0	0/0.26	NP	4: Weak
Dietary factors												
Calcium	HvL	13	1726/351192	0.86 (0.68-1.10)	0.86 (0.74-1.00)	4.0E-02	0.67-1.09	0.62	0	0/1.66	NP	4: Weak
Calcium; dairy	HvL	4	1876/169084	0.86 (0.68-1.10)	0.80 (0.66-0.98)	0.03	0.41-1.59	0.57	34	1/1.50	NP	4: Weak
Cruciferous veg	HvL	11	4306/378489	0.87 (0.74-1.03)	0.89 (0.80-1.00)	0.04	0.71-1.13	0.49	25	1/3.08	NP	4: Weak
Dietary pattern; healthy	HvL	11	3055/395285	0.95 (0.83-1.06)	0.85 (0.74-0.98)	0.02	0.57-1.28	0.27	57	2/0.93	0.24	4: Weak
Dietary pattern; western	HvL	11	3168/328501	1.02 (0.85-1.24)	1.19 (1.00-1.41)	0.04	0.68-2.10	0.87	73	2/0.59	0.06	4: Weak
Dietary protein intake	HvL	10	2354/74464	0.92 (0.82-1.02)	0.92 (0.85-1.00)	3.9E-02	0.83-1.01	0.46	0	0/0.97	NP	4: Weak
Decaffeinated coffee	Ever vs never CC	5	2663/7907	0.71 (0.51-0.99)	0.72 (0.58-0.90)	4.3E-03	0.50-1.04	0.55	0	2/4.03	NP	4: Weak
Dairy Total dairy products	HvL	7	427/90001	1.61 (1.07-2.42)	1.66 (1.19-2.31)	<0.01	NA	NA	NA	1/1.87	NP	4: Weak
Dairy Skim/low fat	HvL	11	728/170327	1.32 (0.97-1.82)	1.35 (1.09-1.68)	<0.01	0.35-5.43	0.21	0	0/1.93	NP	4: Weak
Dairy; whole milk	HvL	9	3590/150503	1.10 (0.75-1.60)	1.25 (1.00-1.56)	0.04	0.67-2.33	0.96	52	2/1.20	0.43	4: Weak
Dairy Lactose	HvL	12	728/170327	1.48 (1.05-2.09)	1.47 (1.17-1.84)	<0.01	0.34-6.29	0.42	0	1/2.64	NP	4: Weak
Egg consumption	HvL	12	3728/632197	1.14 (0.67-1.20)	0.19 (1.00-1.41)	0.04	0.75-1.88	0.56	42	2/2.51	NP	4: Weak
Fat; total fat	HvL	17	7801/760734	1.07 (0.89-1.29)	1.19 (1.04-1.37)	0.01	0.75-1.89	0.57	60	4/1.82	0.09	4: Weak
Fat; trans	HvL	4	2955/152165	1.30 (1.08-1.57)	1.25 (1.08-1.44)	<0.01	0.91-1.71	0.60	0	2/3.29	NP	4: Weak
Meat; Red	HvL	10	5589/358333	1.30 (1.10-1.60)	1.17 (1.02-1.35)	0.03	0.80-1.71	0.95	46	3/7.13	NP	4: Weak
Meat; processed	HvL	7	1018/696100	1.23 (0.92-1.63)	1.26 (1.02-1.56)	3.5E-02	0.31-5.07	0.37	0	0/1.56	NP	4: Weak
Meat; red and processed[κ]	Per 100g/week increment	21	6536/2140286	1.02 (0.98-1.06)	1.01 (1.00-1.04)	3.4E-02	1.00-1.04	0.11	0	0/1.14	NP	4: Weak
Non starchy vegetables	Per 100g/day	6	2053/641079	1.00 (0.93-1.07)	0.94 (0.89-1.00)	4.0E-02	0.82-1.08	0.21	28	1/0.30	0.19	4: Weak
Non herbal tea	HvL	19	734/164882	0.63 (0.40-0.99)	0.69 (0.52-0.93)	1.4E-02	0.11-4.57	0.03	0	1/2.74	NP	4: Weak
Total phytoestrogen intake	HvL	10	4392/298165	1.05 (0.85-1.31)	0.70 (0.56-0.87)	<0.01	0.34-1.43	0.04	72	6/0.78	<1E-100	4: Weak
Isoflavones	HvL	6	2703/164267	0.51 (0.37-0.69)	0.63 (0.46-0.86)	<0.01	0.23-1.70	0.30	68	4/5.99	NP	4: Weak
Daidzein	HvL	4	1239/99776	0.41 (0.29-0.59)	0.60 (0.42-0.86)	<0.01	0.14-2.50	0.33	62	2/ 4.00	NP	4: Weak
Glycitein	HvL	3	959/2501	0.38 (0.37-9.55)	0.55 (0.35-0.86)	<0.01	0.00-86.89	0.22	66	2/ 3.00	NP	4: Weak
Kaempferol	HvL	3	347/66940	0.60 (0.42-0.87)	0.60 (0.42-0.86)	6.0E-03	NA	NA	NA	1/0.99	0.98	4: Weak
Flavones	HvL	4	347/66941	0.69 (0.52-0.92)	0.69 (0.52-0.92)	1.1E-02	NA	NA	NA	1/0.97	0.86	4: Weak
Soybean	HvL	2	754/1906	0.62 (0.45-0.85)	0.48 (0.28-0.84)	<0.01	NA	NA	NA	2/1.93	0.79	4: Weak
Polyphenol; Non asian	HvL	7	3561/232718	1.05 (0.85-1.31)	0.74 (0.58-0.96)	0.02	0.33-1.69	0.13	76	4/0.59	<0.01	4: Weak
Polyphenol; Asian	HvL	3	831/65447	0.45 (0.29-0.59)	0.48 (0.37-0.63)	2.0E-07	0.08-2.87	0.10	0	2/2.97	NP	4: Weak
Tea	HvL	18	8683/708263	1.20 (1.00-1.40)	0.89 (0.80-1.00)	0.04	0.64-1.25	0.01	43	5/8.02	NP	4: Weak

Tea; black	HvL	15	1299/203998	0.63 (0.40-0.99)	0.73 (0.56-0.93)	1.2E-02	0.42-1.24	0.44	15	2/4.71	NP	4: Weak
Tea; green	HvL	4	2580/5991	0.84 (0.64-1.12)	0.63 (0.42-0.93)	0.02	0.11-3.46	0.69	71	2/1.67	0.74	4: Weak
Biochemical markers												
CRP	Middle vs low	7	159/27746	1.52 (1.02-2.28)	1.52 (1.01-2.27)	0.04	NA		NA	1/0.83	0.65	4: Weak
CRP	HvL	7	1898/32580	1.35 (1.06-1.71)	1.35 (1.06-1.71)	0.02	0.67-2.70	0.09	61	4/3.68	0.81	4: Weak
Physical activity												
Recreational	Moderate vs low	20	11765/798786	1.10 (0.95-1.28)	0.91 (0.85-0.99)	0.02	0.71-1.18	0.59	45	4/4.07	1.00	4: Weak

Abbreviations: BMI, body mass index; BMI iya, body mass index in young adulthood; BMI PoMP, body mass index postmenopausal; BMI PrMP, body mass index premenopausal; CC, case control; CRP, c-reactive protein; E+P, estrogen and progesterone; E+E/P, estrogen and estrogen/progesterone; HRT, hormone replacement therapy; HvL, highest versus lowest; IVF, in-vitro fertilisation; NA, not available; np; not pertinent, because the estimated is larger than the observed, and there is no evidence of excess statistical significance based on the assumption made for the plausible effect size; NSAID, non- steroidal anti -inflammatory drugs; OC, ovarian cancer; OCP, oral contraceptive pill; PID, pelvic inflammatory disease; RR, relative risk; SLE, systemic lupus erythematous; WG, weight gain

Key:

*****only meta-analyses meeting at least weak grade of evidence listed

* Number of studies

Relative risk and 95% confidence interval of largest study (smallest standard error) in each meta-analysis

¥ Random effects refer to summary risk ratio (95% confidence interval) using the random-effects model

|| P value of summary random effects estimate

° P-value from the Egger's regression asymmetry test

§ Expected number of statistically significant studies using the point estimate of the largest study (smallest standard error) as the plausible effect size

α Observed/Expected number of statistically significant studies

φP value of the excess statistical significance test

All statistical tests were two-sided

ϥSmall study effect is based on the P-value from the Egger's regression asymmetry test (P>0.1) where the random effects summary estimate was larger compared to the point estimate of the largest study in a meta-analysis

†Based on the p-value (P>0.1) of the excess significance test using the largest study (smallest standard error) in a meta-analysis as the plausible effect size

ι Person years

k Risk of dying from ovarian cancer

**Summary of evidence grading criteria:

Weak	$P<0.05$[ll]
Suggestive	$P<10^{-3}$[ll]; >1,000 cases
Highly suggestive	$P<10^{-6}$[ll]; >1,000 cases; P<0.05 of the largest study in a meta-analysis
Strong	$P<10^{-6}$[ll]; >1,000 cases; P<0.05 of the largest study in a meta-analysis; I^2<50%; no small study effect[q]; prediction interval excludes the null value; no excess significance bias[†]

3.4.8 Sensitivity Analysis

None of the associations that received strong evidence grading were downgraded when both cohort and case-control studies were included (Table 3.7). Two statistically non-significant meta-analyses on talcum powder use were upgraded when case-control studies were included as the evidence on the association between talcum powder use and ovarian cancer risk was derived mainly from case-control studies; ever vs. never use demonstrated highly suggestive and long-term (>10years) vs no use suggestive evidence. Two statistically non-significant meta-analyses on breastfeeding were also upgraded when case-control studies were included; ever versus never use to highly suggestive and breastfeeding >12 months versus never to weak evidence. After the inclusion of case-control studies, the association of parity with ovarian cancer risk was supported with weak evidence although this was previously insignificant.

All six of the risk factors that attained strong evidence (height per 10cm, BMI ≥30kg/m^2 versus normal, all three HRT risk factors and OCP use) remained nominally statistically significant in the main analysis until a 15% credibility ceiling was applied. The three risk factors found to have highly suggestive evidence remained nominally statistically significant until a 10% credibility ceiling was applied (Table 3.8).

I identified 33 associations that had more than one published meta-analysis assessing the same exposure and outcome combination, frequently with several meta-analyses examining different exposure contrasts of the same risk factor (n=41). For most of the duplicate meta-analyses (n=38/41 93%) there was concordance in direction, magnitude and statistical significance of the summary associations between them (Table 3.9).

Table 3.8: Sensitivity analysis using credibility ceilings when the association is non-significant of the 40 studies investigating the risk factors associated with incidence or mortality[$] of ovarian cancer - only cohort studies included.

Author, year	Exposure	Exposure contrast	N[α] studies	Credibility ceiling (%) when association non-significant
Reproductive factors				
Luan 2013	Breastfeeding	Per 5month increase in duration	25	3
Siristatidis 2013	IVF	Ever vs never (ref group general population excluding OC diagnosis <1year post treatment)	6	3
Siristatidis 2013	IVF	Ever vs never ref group IVF pop; total follow up	6	3
Zhou 2017	PID	Ever vs never	14	3
Use of medical or hormonal therapy				
Baandrup 2013	NSAIDS Non aspirin	Ever vs never	16	5
Collaborative Group 2015	HRT- Prospective	Current/recent vs never	12	22
Collaborative Group 2015	HRT- Prospective	Ever vs never	17	19
Collaborative Group 2015	HRT- Prospective	Ever vs never (information on duration of use and time since last use)	14	16
Coll Gr 2008	OCP	Ever vs never	45	22
Hankinson 1992	OCP	Ever vs never	22	12
Shi 2015	HRT	Ever vs never ET only	11	15
Shi 2015	HRT	Ever vs never (continuous E+P)	4	13
Shi 2015	HRT	Ever vs never (sequential E+P)	4	10
Shi 2015	HRT	Ever vs never ET+PT	11	14
Shi 2015	HRT	Ever vs never ET + E/PT	2	9
Zhou 2008	HRT	Current vs ever	5	18
Wen 2019	Metformin	Ever vs never	3	10
Anthropometric indices				
Aune 2015	Anthropometric measure	Height per 10cm	16	25
Aune 2015	Anthropometric measure	BMI per 5kg/m^{2}increase	25	10
Aune 2015	Anthropometric measure	BMI iya per 5kg/m^2 increase	6	13
Aune 2015	Anthropometric measure	Per 5kg weight	4	12
Keum 2015	Anthropometric measure	WG per 5kg increase PoMP, HRT	2	6
Poorolajal 2014	Anthropometric measure	BMI ≥30kg/m^2 vs normal	13	15
Poorolajal 2014	Anthropometric measure	BMI PrMP; ≥30kg/m^2 vs normal	3	8
Poorolajal 2014	Anthropometric measure	BMI PoMP; ≥30kg/m^2 vs normal	5	3
Carcinogens/ environmental irritants				
Reid 2011	Asbestos[$]	Any vs none	16	22

Camargo 2011	Asbestos; Occupational[$]	Total expected vs unexpected	20	14
Camargo 2011	Asbestos; Occupational[$]	High expected vs unexpected	6	7
Medical history				
Bernatsky 2011	SLE	Observed vs expected	5	5
Zhang 2017	Diabetes Mellitus	DM vs no DM	17	9
Li 2019	Endometriosis	Any vs none	25	22
Dietary factors				
Larsson 2006	Dairy Total dairy products	Highest vs lowest	7	9
Larsson 2006	Dairy Skim/low fat	Highest vs lowest	11	12
Larsson 2006	Dairy Lactose	Highest vs lowest	12	13
Song 2017	Calcium	Highest vs lowest	13	6
Wallin 2011	Meat; red and processed[$]	Per 100g/week increment	21	18
Kolahdooz 2010	Meat; processed	Highest vs lowest	7	12
WCRF CUP 2013	Non starchy vegetables	Per 100g/day	6	2
Zhang 2018	Non herbal tea	Highest vs lowest	19	6
Butler 2011	Tea; black	Highest vs lowest	15	6

[α] Number

[$] Risk of dying from ovarian cancer

Key: BMI, body mass index; CC, case-control; DM, diabetes mellitus; ET, estrogen therapy; E+P, estrogen and progesterone; E/PT, combined estrogen/progesterone therapy; g, gram; HRT, hormone replacement therapy; IVF, in vitro fertilisation; iya, in young adulthood; kg, kilogram; m, metre; m^2, metre squared; NSAIDS, non-steroidal anti-inflammatory drugs; OCP, oral contraceptive pill; PID, pelvic inflammatory disease; PoMp, postmenopausal; PrMp, premenopausal; PT, progesterone therapy; ref, reference; SLE, systemic lupus erythematous; WG, weight gain

Table 3.9 Excluded duplicate studies and studies selected in their place to be included in the analysis.

Exposure	Exposure contrast	Author, year	Outcome	N[a] studies	N cohort studies	Summary relative risk & 95% CI[b]	Included	Evidence grade[c]
Alcohol	Any vs none	Rota 2012	OC	27	4	1.00 (0.95-1.05)	No	
Alcohol	Any vs none	Yan-Hong 2015	OC	13	13	1.03 (0.96-1.10)	Yes	NS
Alcohol	Light vs none	Rota 2012	OC	20	3	0.97 (0.92-1.02)	No	
Alcohol	Light vs none	Yan-Hong 2015	OC	13	13	0.96 (0.93-1.00)	Yes	NS
Alcohol	Moderate vs none	Rota 2012	OC	16	3	1.03 (0.96-1.11)	No	
Alcohol	Moderate vs none	Yan-Hong 2015	OC	13	13	1.08 (0.92-1.27)	Yes	NS
Alcohol	High vs none	Rota 2012	OC	4	0	1.09 (0.80-1.50)	No	
Alcohol	High vs none	Yan-Hong 2015	OC	13	13	0.99 (0.88-1.12)	Yes	NS
Aspirin	Any vs none	Bonovas 2005	OC	9	3	0.90 (0.80-1.06)	No	
Aspirin	Ever vs never	Qiao 2018	OC	21	8	0.89 (0.83-0.95)	No	
Aspirin	Ever vs never	Ni 2013	OC	17	3	0.91 (0.82-1.02)	No	
Aspirin	Ever vs never	Zhang 2016	OC	22	8	0.89 (0.83-0.96)	Yes	Weak
Aspirin	Regular vs non regular	Bosetti 2012	OC	11	5	0.91 (0.81-1.01)	No	
Aspirin	Regular vs non regular	Trabert 2014	OC	12	0	0.91 (0.84-0.99)	No	
Aspirin	Regular vs non regular	Murphy 2012	OC	14	4	0.96 (0.85-1.08)	Yes	NS
NSAIDS non aspirin	Regular vs non regular	Bonovas 2005	OC	6	6	0.86 (0.68-1.08)	No	
NSAIDS non aspirin	Ever vs never	Ni 2013	OC	7	3	0.89 (0.74-1.08)	No	
NSAIDS non aspirin	Ever vs never	Baandrup 2013	OC	16	6	0.94 (0.84-1.06)	Yes	Weak
Coffee	Highest vs lowest	Braem 2012	OC	7	7	1.13 (0.89-1.46)	No	
Coffee	Highest vs lowest	Berretta 2018	OC	7	7	1.06 (0.89-1.67)	No	
Coffee	Highest vs lowest	Salari-Moghaddam 2019	OC	13	13	1.10 (0.90-1.35)	Yes	NS
Breastfeeding	Ever vs never	Chowdhury 2015	OC	40	5	0.70 (0.64-0.77)	No	
Breastfeeding	Ever vs never	Feng 2014	OC	19	3	0.66 (0.57-0.76)	No	
Breastfeeding	Ever vs never	Li 2014	OC	40	5	0.69 (0.34-0.76)	Yes	Highly suggestive
Breastfeeding	Ever vs never	Luan 2013	OC	32	5	0.76 (0.69-0.83)	No	
Breastfeeding	Ever vs never, PA	Park 2016	OC	5	0	1.17 (1.02-1.33)	No	
Breastfeeding	Ever vs never	Ip 2007	OC	9	0	0.79 (0.61-0.91)	No	
Breastfeeding	Ever vs never	WCRF CUP 2013	OC	3	3	0.90 (0.75-1.08)	No	

Tubal ligation	Yes vs no, PA	Park 2016	OC	2	0	1.44 (1.33-1.56)	No	
Tubal ligation	Ever vs never	Sieh 2013	OC	13	0	0.71 (0.66-0.77)	No	
Tubal ligation	Ever vs never	Rice 2012	OC	36	6	0.70 (0.64-0.75)	No	
Tubal ligation	Ever vs never	Cibula 2011	OC	21	5	0.69 (0.64-0.75)	No	
Tubal ligation	Ever vs never	Yoon 2016	OC	3	1	0.51 (0.35-0.75)	No	
Tubal ligation	Ever vs never	Wang 2016	OC	25	0	0.70 (0.60-0.81)	Yes	Suggestive
BMI	25-29.9kg/m^2 vs normal	Poorolajal 2014	OC	13	13	1.08 (0.97-1.19)	Yes	NS
BMI	25-29.9kg/m^2 vs normal	Liu 2015	OC	25	12	1.08 (1.02-1.14)	No	
BMI	25-29.9kg/m^2 vs normal, OR	Olsen 2007	OC	14	6	1.16 (1.01-1.32)	No	
BMI	≥30kg/m^2 vs normal	Poorolajal 2014	OC	13	13	1.27 (1.16-1.38)	Yes	Strong
BMI	≥30kg/m^2 vs normal	Liu 2015	OC	25	12	1.23 (1.10-1.39)	No	
BMI	≥30kg/m^2 vs normal, OR	Olsen 2007	OC	16	8	1.30 (1.12-1.50)	No	
BMI	Per 5kg/m2 increase	WCRF Cup 2013	OC	34 (inc pooled study)	22	1.06 (1.02-1.11)	No	
BMI	Per 5kg/m2 increase	Aune 2015	OC	25	25	1.07 (1.04-1.11)	Yes	Suggestive
Cruciferous vegetables	High vs low	Hu 2015	OC	8	4	0.89 (0.81-0.99)	No	
Cruciferous vegetables	High vs low	Han 2014	OC	11	4	0.89 (0.80-1.00)	Yes	Weak
Metformin	Any vs none	Dilokthornsakul 2013	OC	3	0	0.67 (0.44-1.04)	No	
Metformin	Ever vs never	Li 2016	OC	4	1	0.71 (0.54-0.92)	No	
Metformin	Ever vs never	Wen 2019	OC	3	2	0.18 (0.12-0.25)	Yes	Highly suggestive
Diabetes Mellitus	Yes vs no	Wang 2017	OC	14	14	1.19 (1.06-1.34)	No	
Diabetes Mellitus	Yes vs no	Zhang 2017	OCM	17	17	1.32 (1.14-1.52)	Yes	Suggestive
Talc	Exposed vs unexposed	Gross 1995	OC	9	0	1.27 (1.09-1.48)	No	
Talc: on contraceptive diaphragm	Any vs none	Huncharek 2007	OC	9	0	1.00 (0.77-1.29)	Yes	NS
Talc	Ever vs never	Terry 2013	OC	8	0	1.24 (1.15-1.33)	No	
Talc	Ever vs never	Berge 2017	OC	27	3	1.26 (1.17-1.35)	Yes	Highly suggestive
Talc	Ever vs never	Huncharek 2003	OC	16	1	1.33 (1.16-1.45)	No	
Endometriosis	Any vs none; CC/2 arm cohort	Kim 2014	OC	21	3	1.61 (1.01-1.99)	No	
Endometriosis	Any vs none; single arm cohort	Kim 2014	OC	5	5	1.80 (1.28-2.53)	No	
Endometriosis	Ever vs never	Wang 2016	OC	12	0	1.42 (1.28-1.97)	No	
Endometriosis	Any vs none	Li 2019	OC	25	16	1.42 (1.28-1.47)	Yes	Strong
Animal fat	Highest vs lowest	Huncharek 2001	OC	3	1	1.70 (1.43-2.03)	No	

Animal fat	Highest vs lowest	Hou 2015	OC	8	3	1.15 (0.95-1.39)	No	
Animal fat	Highest vs lowest	Qiu 2016	OC	10	5	1.21 (0.99-1.47)	Yes	NS
Dairy fat	Highest vs lowest	Hou 2015	OC	6	5	1.05 (0.92-1.19)	No	
Dairy fat	Highest vs lowest	Qiu 2016	OC	3	2	1.05 (0.92-1.19)	Yes	NS
Monounsaturated fat	Highest vs lowest	Hou 2015	OC	10	3	0.98 (0.84-1.13)	No	
Monounsaturated fat	Highest vs lowest	Qiu 2016	OC	13	5	1.00 (0.90-1.10)	Yes	NS
Plant fat	Highest vs lowest	Hou 2015	OC	9	3	0.93 (0.77-1.13)	No	
Plant fat	Highest vs lowest	Qiu 2016	OC	6	5	0.95 (0.83-1.09)	Yes	
Saturated fat	Highest vs lowest	Qiu 2016	OC	11	6	1.10 (0.99-1.21)	Yes	NS
Saturated fat	Highest vs lowest	Huncharek 2001	OC	3	1	1.20 (1.04-1.29)	No	
Total fat	Highest vs lowest	Huncharek 2001	OC	7	1	1.26 (1.11-1.42)	No	
Total fat	Highest vs lowest	Hou 2015	OC	14	3	1.12 (0.95-1.33)	No	
Total fat	Highest vs lowest	Qiu 2016	OC	17	6	1.19 (1.04-1.37)	Yes	Weak
Fibre	Highest vs lowest	Huang 2018	OC	17	4	0.76 (0.63-0.92)	No	
Fibre	Highest vs lowest	Xu 2018	OC	19	5	0.70 (0.57-0.87)	Yes	NS
Hysterectomy	Yes vs no	Jordan 2013	OC	21	5	0.81 (0.72-0.92)	No	
Hysterectomy	Yes vs no	Rice 2012	OC	36	5	0.74 (0.65-0.84)	No	
Hysterectomy	Yes vs no	Huo 2019	OC	18	0	0.97 (0.83-1.12)	No	
Hysterectomy	Yes vs no	Wang 2016	OC	22	0	0.97 (0.81-1.15)	Yes	NS
Height	Per 10cm	Aune 2015	OC	16	16	1.16 (1.11-1.20)	Yes	Strong
Height	Per 5cm	WCRF CUP 2013	OC	13	13	1.07 (1.05-1.10)	No	
CRP	High vs low	Zeng 2016	OC	7	1	1.91 (1.51-2.40)	No	
CRP	Middle vs low	Zeng 2016	OC	7	1	1.13 (0.96-1.33)	No	
CRP	High vs low	Li 2017	OC	7	1	1.35 (1.06-1.71)	Yes	Weak
CRP	Middle vs low	Li 2017	OC	7	1	1.52 (1.01-2.27)	Yes	Weak
HRT	Ever vs never	Negri 1999	OC	4	0	1.71 (1.30-2.25)	No	
HRT	Ever vs never	Collab Group 2015	OC	17	16	1.20 (1.13-1.28)	Yes	Strong
Dietary α-carotene	Per 600ug/day	WCRF CUP 2013	OC	3	3	1.00 (0.98-1.01)	No	
Dietary α-carotene	Per 600ug/day	Koushik 2006	OC	11	11	1.00 (0.95-1.05)	Yes	NS
Dietary β- Carotene	Per 2500ug/day	WCRF CUP 2013	OC	4	4	0.99 (0.92-1.07)	No	
Dietary β- Carotene	Per 2500ug/day	Koushik 2006	OC	11	11	0.98 (0.93-1.03)	Yes	NS

Dietary β-cryptoxanthin	Per 100ug/day	WCRF CUP 2013	OC	3	3	1.02 (0.90-1.15)	No	
Dietary β-cryptoxanthin	Per 100ug/day	Koushik 2006	OC	11	11	0.99 (0.97-1.02)	Yes	NS
Dietary lycopene	Per 4000ug/day	WCRF CUP 2013	OC	3	3	1.00 (0.93-1.17)	No	
Dietary lycopene	Per 4000ug/day	Koushik 2006	OC	11	11	1.02 (0.98-1.06)	Yes	NS
Tea	Highest vs lowest, HR	Braem 2012	OC	6	6	0.88 (0.71-1.09)	No	
Tea	Any vs none	Zhou 2007	OC	9	2	0.84 (0.66-1.07)	No	
Tea	Highest vs lowest	Zhan 2017	OC	18	7	0.89 (0.80-1.00)	Yes	Weak
Green tea	Highest vs lowest	Butler 2011	OC	4	0	0.63 (0.42-0.93)	Yes	Weak
Green tea	Highest vs lowest	Nagle 2010	OC	2	0	0.58 (0.33-1.01)	No	
Parity	Parous vs Nullip, PA	Park 2016	OC	9	7	0.80 (0.61-1.05)	Yes	NS
Parity	Nullip vs Parous	Negri 1991	OC	3	0	1.40 (1.10-1.70)	No	
Parity	Parous vs Nullip	Whittemore 1993	OC	6	0	0.47 (0.40-0.56)	No	
Acrylamide	Highest vs lowest	Pelucchi 2011	OC	3	2	1.09 (0.76-1.57)	No	
Acrylamide	Highest vs lowest	Pelucchi 2015	OC	4	2	1.12 (0.85-1.47)	Yes	NS
Acrylamide	Per 10ug/day increase	Pelucchi 2011	OC	3	2	1.01 (0.94-1.08)	No	
Acrylamide	Per 10ug/day increase	Pelucchi 2015	OC	4	2	1.02 (0.96-1.09)	Yes	NS
Sedentary behaviour	Highest vs lowest	Shen 2014	OC	2	2	1.26 (0.87-1.82)	No	
Sedentary behaviour	Highest vs lowest	Schmid 2014	OC	5	2	1.22 (0.97-1.70)	Yes	NS
OCP	Ever vs never	Coll Gr 2008	OC	45	13	0.74 (0.70-0.78)	Yes	Strong
OCP	Ever vs never	Hankinson 1992	OC	20	3	0.72 (0.54-0.97)	No	
OCP	Never vs Ever, PA	Park 2016	OC	2	1	1.87 (0.89-3.94)	No	
OCP	Ever vs never	Zhong 2015	OC	3	3	0.58 (0.35-0.94)	No	

Abbreviations: BMI, body mass index; CC, case control; CI, confidence interval; Coll, collaborative; CRP, C reactive protein; Gr, group; HR, hazard ratio; HRT, hormone replacement therapy; inc, including; kg, kilogram; m^2, metre squared; NS, non-significant; NSAID, non-steroidal anti-inflammatory drug; OCP, oral contraceptive pill; OR, odds ratio; PA, pooled analysis; ug, microgram; yoa, years of age; WCRF, World Cancer Research Fund

Key:

[a] Number of studies

[b] Summary relative risk of random effects model including all study types

[c] Summary of evidence grading criteria:

[d] P value of summary random effects estimate

[e] Small study effect is based on the P-value from the Egger's regression asymmetry test (P>0.1) where the random effects summary estimate was larger compared to the point estimate of the largest study in a meta-analysis

[f] Based on the p-value (P>0.1) of the excess significance test using the largest study (smallest standard error) in a meta-analysis as the plausible effect size.

Weak	$P<0.05$[d]
Suggestive	$P<10^{-3}$[d]; >1,000 cases
Highly suggestive	$P<10^{-6}$[d]; >1,000 cases; P<0.05 of the largest study in a meta-analysis
Strong	$P<10^{-6}$[d]; >1,000 cases; P<0.05 of the largest study in a meta-analysis; $I^2<50\%$; no small study effect[e]; prediction interval excludes the null value; no excess significance bias[f]

3.5 DISCUSSION

Main Findings and Interpretation:

This umbrella review explored the associations of 55 non-genetic risk factors with the risk of developing or dying from ovarian cancer and included 212 meta-analyses. Only six associations were graded with strong evidence. Greater height and BMI and the use of HRT were associated with a higher risk of developing ovarian cancer whilst the use of OCP was associated with lower disease risk. Further positive associations between HRT use as well as endometriosis received a highly suggestive evidence grade, whilst metformin use had highly suggestive evidence for lower disease risk.

The finding for height as a risk factor for ovarian cancer incidence is in agreement with the WCRF CUP report, which judged the evidence as 'convincingly strong'[12]. Two mechanisms have been postulated: increased total number of stem cells with subsequent increased chance of DNA mutations during cell division and/or raised insulin-like growth factor 1 (IGF1) (major determinant of height) and its link to a number of solid tumours including ovarian tumours[233,234]. However, a recent Mendelian randomisation study showed evidence for an association of height only with clear cell carcinoma (OR 1.36; 95%CI 1.15–1.61) but not with other subtypes[235]. Most observational studies have not performed analyses by histological subtypes of ovarian cancer; thus, I was unable to assess these associations in the current umbrella review. Future large epidemiological studies or consortia thereof should address potential differential associations by ovarian cancer subtypes.

The strong evidence on the association of obesity with ovarian cancer risk is also in line with previously published umbrella reviews and the World Health Organisation (WHO)/International Agency for Research on Cancer (IARC) World Cancer Report 2020[16,123,124]. This was also supported by a Mendelian study which suggested evidence of a link between obesity and invasive epithelial ovarian cancer[235]. This study also presented evidence suggesting that BMI is differentially associated by histological subtypes of ovarian cancer. Positive associations of BMI with high grade serous and endometrioid carcinomas were reported, but associations for low grade serous, mucinous and clear cell carcinomas had wide confidence intervals including the null value. A pooled analysis of observational studies in the Ovarian Cancer Association Consortium reported an increased risk of high BMI with total ovarian cancer, which was more pronounced for borderline serous, invasive endometrioid and invasive mucinous tumours[236]. Whilst this analysis examined the association between BMI and ovarian cancer throughout different life stages i.e., in young adulthood, pre- and post-menopause, the included studies did not contain subanalyses by histology. Several pathways have been proposed linking obesity with cancer including the IGF-axis signalling and hyperinsulinaemia, alterations in sex hormone metabolism, and subclinical chronic low-grade inflammation with alterations in adipokine pathophysiology[237,238]. Both

insulin and leptin have been shown to promote the growth of cancer cells, whilst leptin also functions as an inflammatory cytokine.

Several hormonal and reproductive factors have been previously associated with ovarian cancer. The findings in this umbrella review confirm strong evidence to support higher risk with HRT and lower risk with OCP use. Breastfeeding reported highly suggestive evidence for lower risk, whilst nulliparity showed weak evidence of a higher risk when both cohort and case-control studies were included[188,196].

The impact of HRT use was confirmed in all nine meta-analyses using different HRT formulations (e.g., oestrogen alone, oestrogen and progesterone continuously, oestrogen and progesterone sequentially) and in the WHO/IARC World Cancer Report (WCR) 2020[16]. No Mendelian randomisation studies on HRT use exist due to the lack of robust genetic variants available to serve as proxies. The WHO/IARC WCR 2020 also agreed with this analysis for an inverse association for OCP use (RR 0.74, 95% CI 0.69-0.80)[16].

The mechanisms that underpin these associations are complex. The concept of 'incessant ovulation' supports the notion that recurrent minor trauma caused to the ovarian epithelial surface as a result of ovulation increases the risk of malignant transformation. This process may explain the reduction of risk with OCP, breastfeeding and parity as they are associated with fewer ovulatory cycles[239-243]. Furthermore, it has been suggested that prolonging the ovarian epithelial surface proliferation stimulated by oestrogen exposure e.g., in HRT could result in ovarian oncogenesis. Indeed, 60% of ovarian tumours are oestrogen-receptor positive[244]. Several mechanisms have been suggested linking oestrogens with ovarian cancer. These include i) a direct transcriptional effect of oestrogens on target genes via activation of the oestrogen receptor-α leading to the transcription of several genes which stimulate cell proliferation and thus an increased risk of mutations, ii) promotion of tumour progression through influencing of signalling pathways leading to the activation of epidermal growth factor receptor, and iii) generation of free radicals from the metabolic activation of catechol oestrogens causing neoplastic transformation of cells[245]. Conversely, progesterone exposure may counteract these effects. Progesterone receptors are also present in normal ovarian epithelium, and it has been suggested that progestins may counteract the proliferative effect of oestrogen by prompting apoptosis within the ovary[246]. This plausible explanation has been reflected both in this study where oestrogen-only HRT was supported by highly suggestive evidence, whilst mixed preparations demonstrated weak evidence, and also in observational studies where oestrogen plus progestin HRT preparations display a slightly weaker effect than oestrogen-only preparations[247,248]. OCP studies have also consistently demonstrated a reduction in ovarian cancer risk with cyclical administration of progestins[240,249-251].

The positive association of endometriosis with ovarian cancer was supported by highly suggestive evidence when using only cohort studies, and it was upgraded to strong evidence when case-control studies were included. Endometriosis has been primarily associated with clear cell (OR 3.05, 95% CI 2.43–3.84) and endometrioid (OR 2.04, 95% CI 1.67–2.48) ovarian cancer in the published literature, although the studies included in this analysis did not provide data by histological subtype[252]. The Mendelian randomization study by Yarmolisky et al. confirmed these findings (invasive epithelial ovarian cancer, OR 1.10, 95% CI 1.06-1.15; endometrioid, OR 1.14, 95% CI 1.04-1.24; clear cell, OR 1.49, 95% CI 1.29-1.73)[235]. Several mechanistic theories have been postulated linking endometriosis to these endometriosis-associated ovarian cancers (EAOC). Loss of tumour suppressor genes *PTEN* and *ARID1A*, caused by oxidative stress from iron-rich endometriotic cysts have been found in both EAOC tumours and adjacent endometriotic tissue[253]. 'Retrograde menstruation' has also been extensively discussed whereby endometrial cells transplanted to the pelvic peritoneum and ovaries undergo neoplastic change as a result of inflammatory processes leading initially to atypical endometriosis and subsequently neoplastic change[254]. This concept is further supported by my finding of suggestive evidence that tubal ligation reduces the risk of ovarian cancer (OR 0.70, 95% CI 0.60-0.81).

This umbrella review revealed suggestive evidence that diabetes increases the risk of ovarian cancer incidence, which is in line with a previously published assessment of the literature by Tsilidis and colleagues (RR 1.17, 95%CI 1.02- 1.34)[159,221]. Furthermore, ever use of metformin was found to have highly suggestive evidence for a lower ovarian cancer risk (RR 0.18, 95% CI 0.12-0.25)[213]. Several mechanisms linking diabetes and oncogenesis have been described. It has been suggested that insulin resistance could induce the expression of cytokines which have tumour promoting effects, elevated blood glucose is associated with increased vascular endothelial growth factor thus promoting angiogenesis and aiding tumour proliferation and hyperinsulinaemia could increase the levels of bioavailable oestrogen through a reduction in sex hormone binding globulin (SHBG) levels[255].

There has been intense debate as to whether talcum powder use increases the risk of ovarian cancer[256]. The multinational cosmetics company Johnson & Johnson are currently appealing a $4.69bn verdict against it from July 2018 in Missouri, United States of America (USA), towards 22 plaintiffs who claim that the use of their talcum powder led to ovarian cancer due to contamination with asbestos and heavy metals. A further ruling in April 2020 in New Jersey, USA has dictated that a new lawsuit against the company for the same claim may proceed after an appeal against inclusion of certain expert testimonies. The main body of evidence derives from case-control studies. The meta-analysis comparing ever vs never use included 27 studies, of which only three were cohort studies. Although none of the cohort studies reported significant results, the evidence grading was highly suggestive when case-control studies were included, and this would have been graded as strong if the 95% PI did not just include unity. Case-control studies are susceptible to recall bias which has driven the main analysis to include

cohort studies only. Therefore, the evidence of this association was not considered as robust, which is supported by a recent pooled analysis of four prospective cohort studies showing no association[257].

3.6 STRENGTHS AND LIMITATIONS

In this umbrella review, I have presented the most comprehensive critical appraisal of published observational associations between non-genetic risk factors and the risk of developing or dying from ovarian cancer to date. A wide array of statistical tests and sensitivity analysis were used to assess evidence strength and validity to create a systematic and transparent method of evidence appraisal of the literature.

Several limitations and caveats should also be considered when interpreting these findings. Only published studies were included in this work as is standard practice. To ensure no potentially relevant articles were missed, I conducted a thorough hand search of my personal database and reference list of included articles which not reveal any further unpublished relevant studies at the time of the literature search. Although some individual studies may have been missed in the literature search of the original meta-analyses, this is not expected to have altered my inference as duplicate meta-analyses on the same exposure-outcome pairs gave comparable results. The quality of the original studies included in the meta-analyses was also not assessed and was beyond the scope of this umbrella review, but I did assess the quality of the meta-analyses. Any exposure which did not demonstrate any statistically significant results was not shown in this work for brevity but have been noted. Furthermore, the statistical tests used can 'hint' the presence of possible bias but are unable to confirm its definitive presence and source. The analysis was further restricted by the inability to explore the impact of risk factors separately on pre- and post-menopausal women, for different histological subtypes (e.g., serous or endometrioid), the preparation and length of use of HRT or OCP due to the lack of study-specific data provided by the published meta-analyses[153].

3.7 CONCLUSIONS

This umbrella review provides a comprehensive summary of the body of published systematic reviews and meta-analyses examining risk factors and the incidence and mortality from ovarian cancer. The positive associations of height, obesity and use of HRT with ovarian cancer risk, and the inverse association of OCP use with disease risk were supported by strong evidence. There was further highly suggestive evidence that endometriosis increases disease risk, whilst metformin reduces it. Although further exposures may also be associated with ovarian cancer, the evidence is less certain and further research is required.

The description of genuine risk factors found to alter ovarian cancer incidence and mortality is of great interest to the scientific community, clinicians and the public given the dearth of preventative strategies for ovarian cancer. Information on hormonal modifiable risk factors such as HRT and OCP use can be used in the counselling of women considering their use and how this may positively or inversely affect their risk, whilst further research on the mechanisms that promote reduction of risk in exposures such as the oral contraceptive pill is required. Drugs like metformin or others working on similar metabolic pathways may be further used and are possible chemopreventative targets. More importantly, and in the absence of efficient screening and prevention, evidence on the strength of previously proposed associations may enhance awareness of individuals at higher risk of ovarian cancer and promote prevention or early detection.

3.8 STATEMENT OF CONTRIBUTION

Statistical analysis of extracted data with Stata software (version 15) was conducted by Ilkka Kalliala.

CHAPTER FOUR: INVESTIGATING THE PRESENCE OF A GENUINE BACTERIAL SIGNATURE IN THE FEMALE UPPER GENITAL TRACT AND THE PRESENCE OF A BACTERIAL CONTINUUM

4.1 INTRODUCTION

The advent of next generation sequencing (NGS) has led to rapid progression in our understanding of our microbial cohabitants and their role in our homeostasis, specifically in relation to health and disease. Before NGS, and specifically 16S rRNA gene sequencing, investigators were reliant on culturing of bacterial samples or microscopy. 16S rRNA sequencing enabled fuller and more accurate profiling of bacterial communities resulting in huge data acquisition. Extensive examination of several microbiomes within the human body has since occurred including, for example, the gastrointestinal system[258,259], skin[260], oral[261] and nasopharynx[262] and vagina[100].

Subsequently, the microbial communities in the lower genital tract (LGT) have been well described in both health and disease states e.g., pelvic inflammatory disease, bacterial vaginosis, cervical dysplasia[100,101,104,263,264], in part due to the accessibility and relatively high biomass of the LGT. Further efforts to identify and characterise microbiomes from 'low biomass' sites such as the upper genital tract (UGT), however, have been dogged with complications related to sample contamination and low microbial density. Early experiments which were naïve to these problems did not control adequately for potential contaminants. These contaminants could arise from several sources including environmental, biological and material e.g., sequencing kit. Genetic material could be translocated from the kit being used for the extraction of the DNA (the 'kitome'), or from the sequencing machines themselves, with leftover genetic material 'lane hopping' from previous experiments[265,266].

More recent studies of the UGT have attempted to control for these low biomass issues. The long-held belief that the UGT was a sterile environment was initially questioned almost thirty years ago[267], but it was not until recent years, assisted by 16S rRNA sequencing, that a microbial signature was discovered in the endometrial cavity[107]. Subsequent studies investigating the presence of microbes in the UGT have trialed various sources of samples e.g., swabs, tissue, peritoneal fluid, in an effort to determine if a bacterial signature above the background level was present[106,143,268,269]. Most of these have focused on the endometrial cavity, with few interrogating further upstream to the ovaries and fallopian tubes[33,34]. A handful of these studies on the endometrial cavity have also attempted to elucidate if there is a common link, or continuum, of microbes which are identified throughout the genital tract[106,268]. Translationally, if reliable continuums for both benign and malignant conditions were discovered, this would enable conditions of the UGT prone to dysbiosis to be identified from the more easily accessible LGT thereby revolutionizing clinical diagnostic pathways.

Of the few studies concentrating on the ovaries and fallopian tubes, most have examined the upper genital tract microbes, with only Nené et al. and Qin et al. assessing the role of the cervicovaginal microbiome in ovarian cancer[33–35,38]. Two further studies have very recently assessed the genital tract as a whole in relation to ovarian cancer[37,38].

In light of the now established role of serous tubal intraepithelial carcinomas (STIC) within the fallopian tubes as a precursor to ovarian cancer, and the link between inflammatory conditions of the upper genital tract and ovarian cancer, further investigation into the microbial landscape and continuum of the genital tract in relation to ovarian cancer is warranted. Well planned experiments to overcome the contamination hurdles which low biomass sites can encounter to assess the microbial milieu of the entire genital tract in ovarian cancer could help to further unravel some of the causative mechanisms and potential future diagnostic pathways of this heterogenous group of conditions.

4.2 AIMS

1. Investigate the presence of bacteria in the female upper genital tract taking careful consideration of potential environmental and reagent contaminants

2. Investigate if there is a continuum, or evidence of potential migration of bacteria from the lower genital tract to the upper genital tract in female patients

4.3 RESULTS

4.3.1 Study populations characteristics

A total of 31 women were prospectively recruited into this study. Nineteen women had ovarian or tubal cancer, and 12 women had benign conditions. Of those, the full collection of swab samples from throughout the genital tract was collected from 23 patients; 15 of whom had malignant disease, and eight who had benign conditions.

Patients were classified according to tissue histology. Due to the small cohort size, those with ovarian cancer were not subclassified by grade or stage. The average age of the cohort was 60 ($P > 0.99$) (Table 4.1). There were proportionally more Caucasian women in the malignancy cohort than the benign cohort ($P= 0.051$). Woman with a malignancy were less likely to have had pelvic inflammatory disease ($P= 0.005$), but no differences between the two groups were noted in parity ($P= 0.243$) menopausal status ($P= 0.627$), HRT ($P= 0.722$) or OCP use ($P= 0.574$, or BMI ($P= 0.432$). Although more women with cancer ($n=3$) required antimicrobials than in the benign group ($n=1$) there was no significant difference ($P= 0.666$). Two of the women with cancer required antimicrobials for suspected genital tract

infections, with the remaining woman requiring a topical antifungal for a cutaneous Candida infection. The patient with benign disease was prescribed a topical antifungal for a vaginal Candida infection.

Including both control (n=30) and biological samples (samples from the patients e.g., vaginal, cervical etc.) (n=186), a total of 8,408,869 reads from 216 samples with a mean of 1,515 reads per sample were captured. A total of 4,924 (1,565) taxa were identified after removal of singletons and rare OTUs.

4.3.2 Contamination considerations in microbiome research

4.3.2.1 Contaminated samples

Two hundred and sixteen swab samples including patient and control samples were processed. Thirteen control samples were removed as they were not relevant to the cohort of samples analysed as they were from recruitment days of patients not included in this cohort e.g., patients who were excluded due to incomplete metadata and samples etc. Two additional control samples were removed after initial review due to patterns of sequencing data which indicated likely contamination. When reviewing the sequencing data of the controls, it was noted that the two negative controls displayed similarly considerable amounts of bacteria in the same frequency and prevalence as the study samples which had been extracted in the same group indicating likely contamination of the controls during the DNA extraction. A further six samples, including five patient samples and one control sample, were removed due to very abnormal and unexpected taxonomy which indicated both cross contamination and external DNA contamination (Figure 4.1). Upon further investigation, these six samples were found to have been acquired from a patient who underwent a bowel resection intra-operatively, leading to sample DNA contamination from bowel contents, with evidence of cross contamination also noted. A total of 197 samples were therefore included for subsequent analyses.

Table 4.1 Patient characteristics at baseline by histology status.

Characteristics	Benign (n=12)	Cancer (n=19)	Total (n=31)	P value* (Benign vs cancer)
Age, years				
Mean(SD, range)	60 (14.1, 52-89)	60 (13.52, 41-90)	60 (13.52, 41-90)	>0.999 #
Ethnicity, n/N (%)				0.051
Caucasian	6/12 (50)	14/19 (73.7)	20/31 (64.5)	
Asian	0/12	3/19 (15.8)	3/31 (9.7)	
Black	4/12 (33.3)	2/19 (10.5)	6/31 (19.4)	
Unknown	2/12 (16.7)	0/19 (0)	2/31 (6.5)	
Parity, n/N (%)				0.243
Nulliparous	2/12 (16.7)	4/19 (21)	6/31 (19.4)	
Parous	8/12 (66.7)	15/19 (78.9)	23/31 (74.2)	
Unknown	2/12 (16.7)	0/19 (0)	2/31 (6.5)	
Menopausal status, n/N (%)				0.627
Premenopausal	4/12 (33.3)	3/19 (15.8)	7/31 (22.6)	
Postmenopausal	8/12 (66.7)	15/19 (78.9)	23/31 (74.2)	
Unknown	0/19 (0)	1/19 (5.3)	1/31 (3.2)	
HRT use, n/N (%)				0.722
Never	10/12 (83.3)	17/19 (89.5)	29/31 (87)	
Ever	1/12 (8.3)	2/19 (10.5)	3/31 (9.7)	
Unknown	1/12 (8.3)	0/19	1/31 (3.2)	
OCP use, n/N (%)				0.574
Never	5/12 (41.7)	7/19 (36.8)	12/31 (38.7)	
Ever	5/12 (41.7)	11/19 (57.9)	16/31 (51.6)	
Unknown	2/12 (16.6)	1/19 (5.3)	3/31 (9.7)	
PID, n/N (%)				0.005
Never	7/12 (58.3)	19/19 (100)	26/31 (83.9)	
Ever	2/12 (16.7)	0/19 (0)	2/31 (6.5)	
Unknown	3/12 (25)	0/19 (0)	3/31 (9.7)	
BMI, n/N (%)				0.432
Non-obese (BMI <30kg/m2)	6/12 (50)	13/19 (68.4)	19/31 (61.3)	
Obese (BMI ≥30kg/m2)	6/12 (50)	5/19 (26.3)	11/31 (57.8)	
Unknown	0/12	1/19 (5.3)	1/31 (3.2)	

BMI, body mass index; HRT, hormone replacement therapy; OCR, oral contraceptive pill; PID, pelvic inflammatory disease; SD, standard deviation.

*Calculated using Fisher's exact test. A p-value of less than 0.05 indicates a significant difference in the distribution of the demographic of interest (e.g., parity), according to histological status.

Calculated using a t-test. A p-value of less than 0.05 indicates a significant difference in the distribution of age according to histological status.

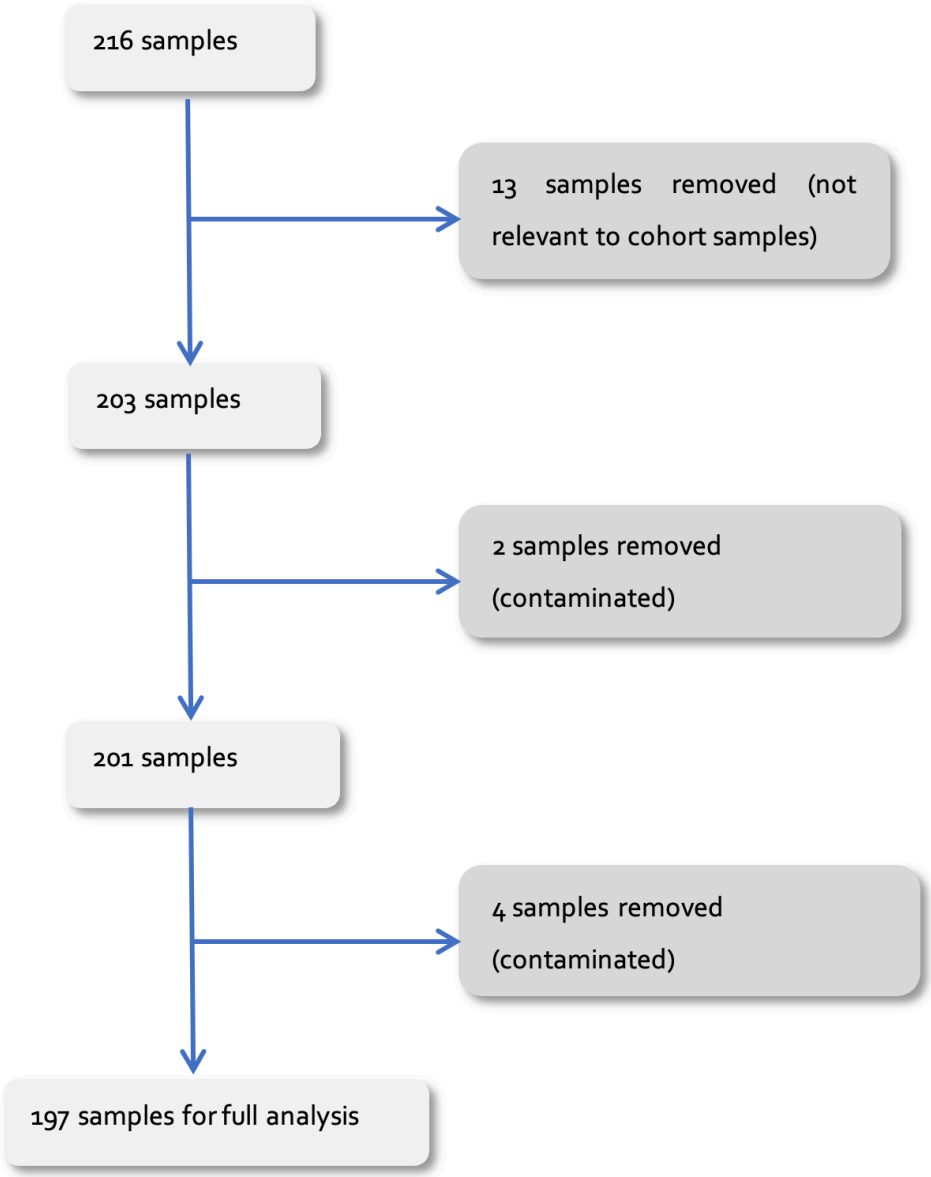


Figure 4.1 Flow diagram of sample selection

Table 4.2 Repeating patterns of sequence reads across all samples indicates carry over of species DNA from the first samples and subsequent contamination of samples further along the processing line.

Species	Cervix	Vagina	Ovary	Fallopian tube	Endometrium	Negative control
	E20	E21	E22	E23	E24	E25
Lactobacillus_iners	5	35	55	230	286	9505
Gardnerella_vaginalis	0	0	6	0	1	1
Lactobacillus_crispatus	212	321	272	226	183	13
Lactobacillus_gasseri	27530	26527	1857	1833	1486	714
Lactobacillus_acidophilus	1	5	21	65	212	20971
Lactobacillus_fornicalis	0	1	2	0	8	687
Lactobacillus_leichmannii	0	3	0	0	0	0
Prevotella_copri	0	0	0	0	0	0
Alloscardovia_omnicolens	440	334	0	0	0	2
Atopobium_vaginae	0	0	0	0	1	0
Bifidobacterium_longum	730	801	657	506	724	0
Streptococcus_agalactiae	0	3	0	0	1	0
Bacteroides_vulgatus	242	274	516	732	607	3
Pseudocitrobacter_anthropi	46	132	397	380	264	0
Anaerococcus_lactolyticus	0	0	0	0	0	0
Prevotella_unclassified	616	913	250	304	337	0
Faecalibacterium_unclassified	8069	10016	7400	11124	10717	352
Finegoldia_magna	4	50	0	0	5	0
Streptococcus_anginosus	3	15	7	4	5	0
Prevotella_timonensis	1	4	0	0	0	0
Anaerococcus_unclassified	0	1	4	7	0	23
Prevotella_bivia	0	0	0	1	0	0
Bifidobacterium_dentium	3031	6159	4	15	13	642
Peptoniphilus_harei	3	25	0	16	0	61
Bacteroides_gallinaceum	9785	11903	8592	6971	10745	350
Bacteroides_massiliensis	0	0	0	2	0	1
Faecalibacterium_prausnitzii	4020	4993	3262	4893	4396	197
Ureaplasma_parvum	448	491	1	0	0	0
Bacteroides_dorei	6941	7919	5910	5216	7437	62
Oscillibacter_unclassified	3665	5194	3917	4301	4180	130
Propionibacterium_acnes	0	1	3	11	13	146
Actinotignum_schaalii	0	0	0	0	0	0
Streptococcus_mitis	36	50	43	57	50	0
Bifidobacterium_breve	0	6	0	0	0	0

4.3.2.2 Process used to identify and manage background contamination in low biological samples

4.3.2.2.1 Manual assessment of contamination

For manual assessment and identification of reads likely derived from non-patient sources, a ratio of the average number of reads per sample to average number of reads per control was calculated. Any OTUs with a ratio less than two, indicating that the control samples had similar amounts of biomass as the patient samples, were removed leaving 3871 OTUs which represented the remaining OTUs with a value above background control levels.

4.3.2.2.2 Identification and removal of contamination sequences using the Decontam statistical package

Shortly after data analysis commenced, an open-source R package, decontam, was released which aims to address the issue of DNA contamination identification in sequencing data sets.[148] Decontam implements a statistical classification procedure that identifies contaminants in microbiome data using the same widely recognised patterns of frequency and prevalence. Decontam is particularly useful for accurate classification of 'true' sequences in undefined microbial communities with low biomasses when comparing against negative control samples. Using a threshold of $P < 0.5$, i.e., a more stringent cut off (accepting the risk that 'false' contaminants may be removed), 1007 OTUs were removed. $P<0.1$ is less stringent (accepting the risk that 'true' contaminants may not be removed), resulting in 363 OTUs being removed (Figure 4.2). When comparing the two decontam methods, and subsequently the decontam and manual results, the data resulting from the $P < 0.5$ threshold was utilised. The $P<0.5$ threshold incorporated all the potential contaminants identified by the $P<0.1$ threshold, and approximately 80% of those identified by the manual method. As decontam is a more robust and objective approach than manual removal, further data analyses was performed on 'true' sequences (n=1601 OTUs) identified using this method.

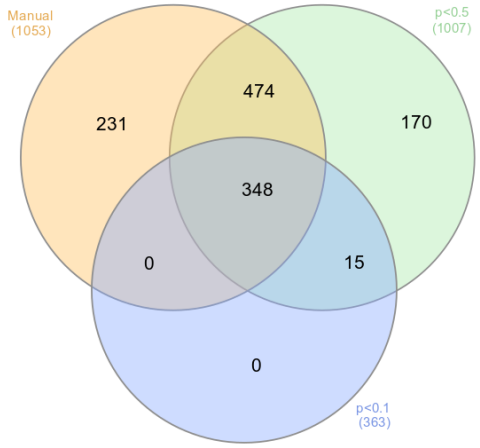


Figure 4.2 Venn diagram demonstrating the results of the three different thresholds for identifying suspected contaminated DNA sequences. All of the P<0.1 results were incorporated into the P <0.5 group. 78% from the manual suspected contaminant DNA sequences were included in the P <0.5 group.

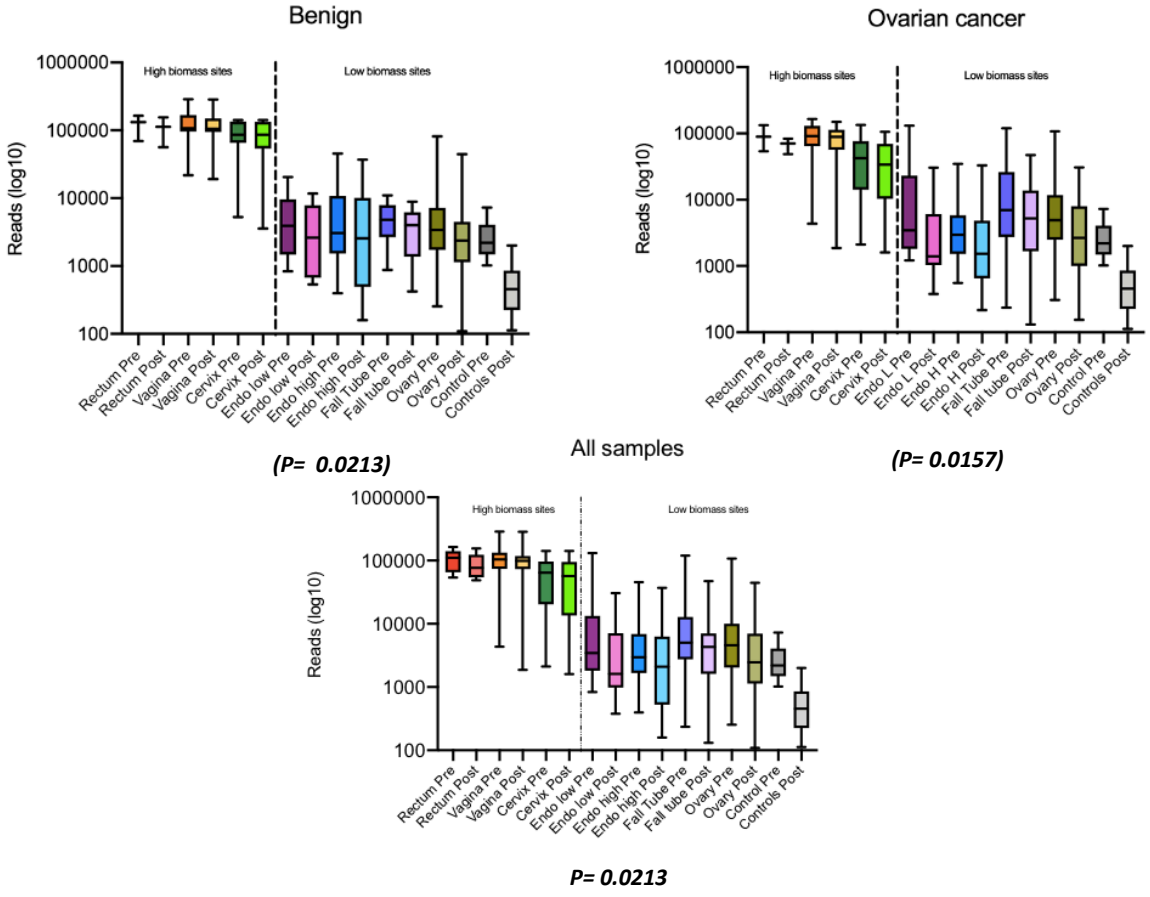


Figure 4.3 Difference in read counts pre and post application of decontam package per site in the benign (P= 0.1704 corrected, P= 0.0213 uncorrected) and cancer (P= 0.1259 corrected, P= 0.0157 uncorrected for control samples) groups, and all samples (P= 0.1704 corrected, P= 0.0213 uncorrected). All P values of the controls calculated using Kruskall Wallis with Dunn's post hoc test.

*'Pre' = before application of the decontam package, and 'post' is after the package.

A statistically significant difference in the total number of read counts before and after application of the decontam package was detected at all sites in the whole patient cohort and the negative controls (Figure 4.3, P= 0.0213). A difference was also observed at all samples sites from the benign group (Figure 4.3, P= 0.0126) and ovarian cancer group (Figure 4.3, P= 0.0157). The relative percentage of contaminant reads versus 'true' reads for each sample type was also examined using a stacked bar chart (Figure 4.4). The negative controls and low biomass samples (ovary, fallopian tube and endometrium) were shown to harbour the highest proportion of contaminant reads.

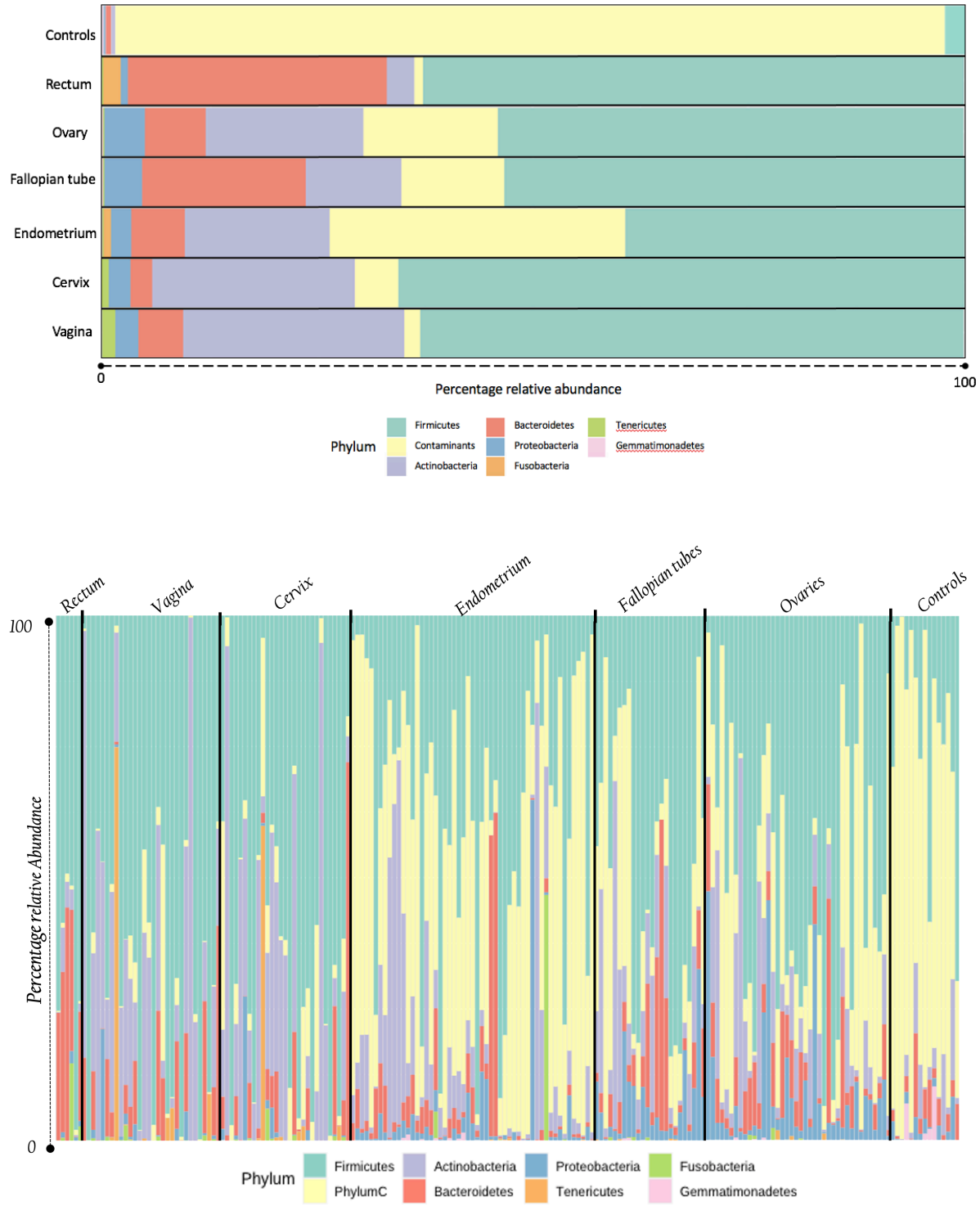


Figure 4.4. Depiction of the proportion of contaminants (yellow) in each collection of samples by anatomical site and control group (a) and individual samples per site (b). The figure illustrates how the upper genital tract samples had a disproportionately higher volume of contaminants than the lower genital tract.

4.3.2.3 Bacterial taxonomic composition at the higher and lower endometrial sites is comparable

Samples were taken from sites at the endocervical junction and at the fundus of the endometrium. A two-group comparison of the swabs from the different sites was performed indicating that relative abundance of bacterial taxa was highly comparable between the upper and lower groups (R^2=0.963 from phylum to genus level, Figure 4.5a, R^2=0.956 at species level, Figure 4.5b).

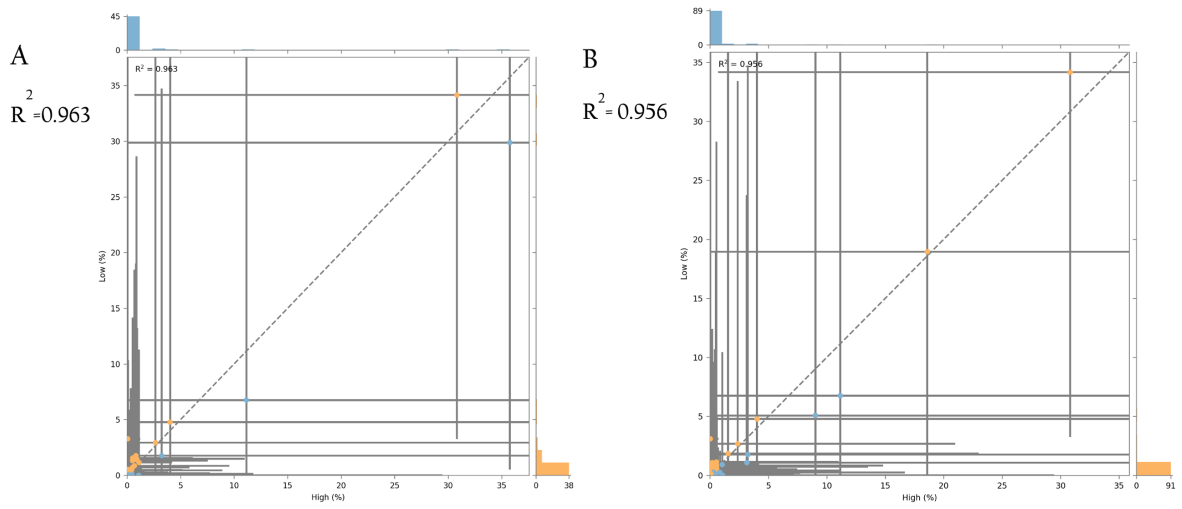


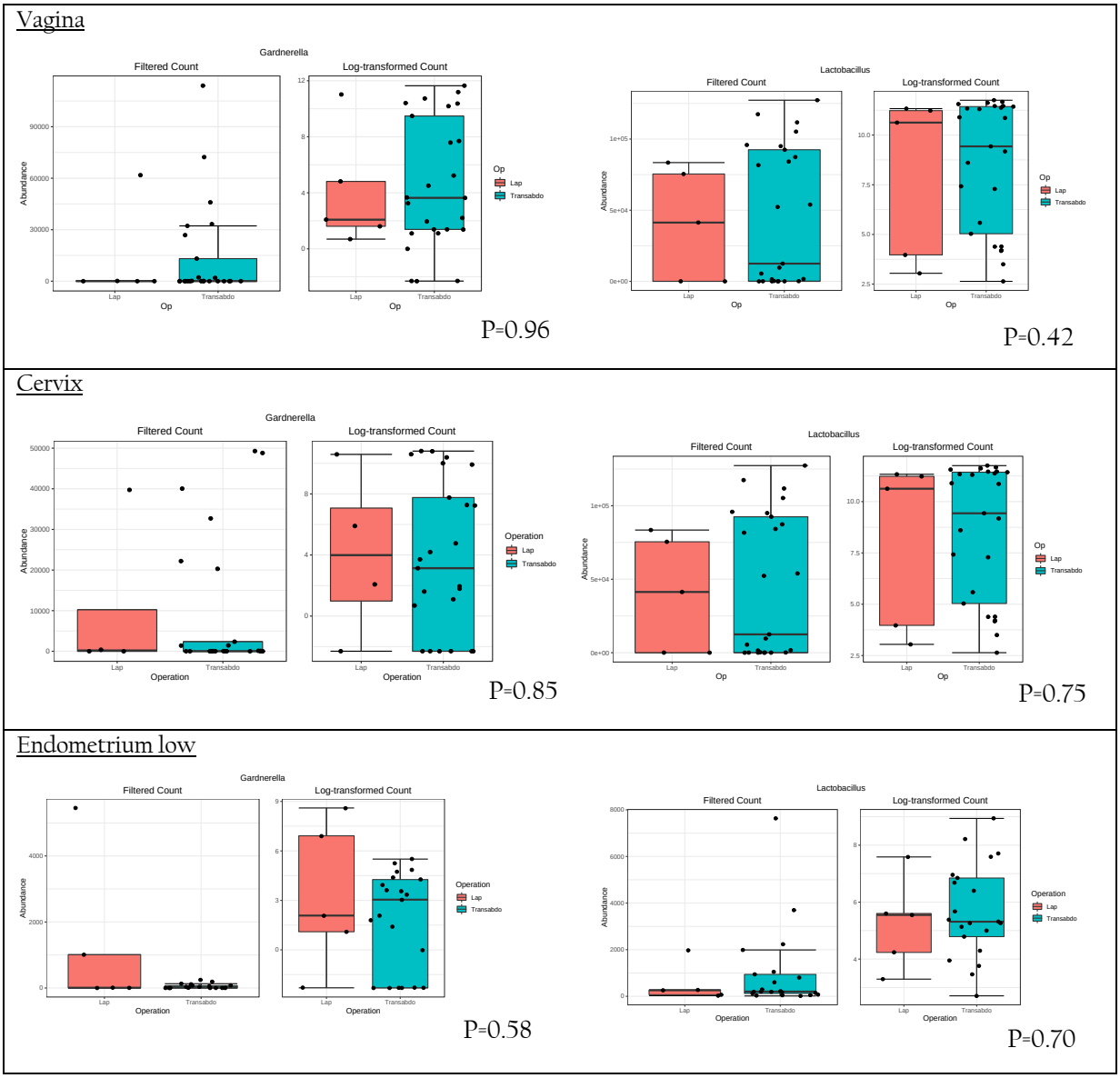
Figure 4.5. Correlation analysis of upper and lower endometrial samples at A. genera level and B. species level. The upper endometrial samples taken from the fundus of the uterus were compared with the samples taken from the endocervical junction at the inferior aspect of the uterus and were assessed at both genera (R^2=0.963) and species (R^2= 0.956) level. Both results demonstrated high levels of similarity between the two sites.

4.3.2.4 Bacterial composition is comparable between samples taken from specimens removed transabdominally and laparoscopically.

Ovarian cancer surgery is predominantly performed as an 'open' procedure, where a midline incision is made vertically in the abdomen, enabling access to the entire abdominopelvic cavity, and to the thorax, should it be required. This also allows the diseased organs and tissue to be removed 'en bloc' through the abdominal incision. Many other gynaecological operations, including several oncological procedures, are performed laparoscopically, or via 'keyhole' surgery. In total laparoscopic hysterectomies and bilateral salpingo-oophrectomies, where the cervix, uterus, and both fallopian tubes and ovaries are removed, the specimen is removed via an incision made below the level of the cervix at the apex of the vagina. Samples for this study were taken from specimens removed via both abdominal and laparoscopic procedures. There was potential for contaminants to be picked up from

the vagina during laparoscopic removal. The bacterial composition from each method was thus compared to identify if the composition differed depending on how the specimen was removed. Twenty-four specimens removed transabdominally were compared with five specimens removed laparoscopically. Univariate analysis indicated no difference in the bacterial composition between the samples removed laparoscopically or transabdominally at any of the anatomical sites within the genital tract when $P<0.05$ was applied at genus level (Figure 4.6). Gardnerella and Lactobacillus were reviewed specifically as they typically represent the most abundant bacteria in the vagina.

Should there have been significant contamination of the specimens removed transvaginally during the laparoscopic procedure, higher read counts of Lactobacillus and Gardnerella would have been expected.



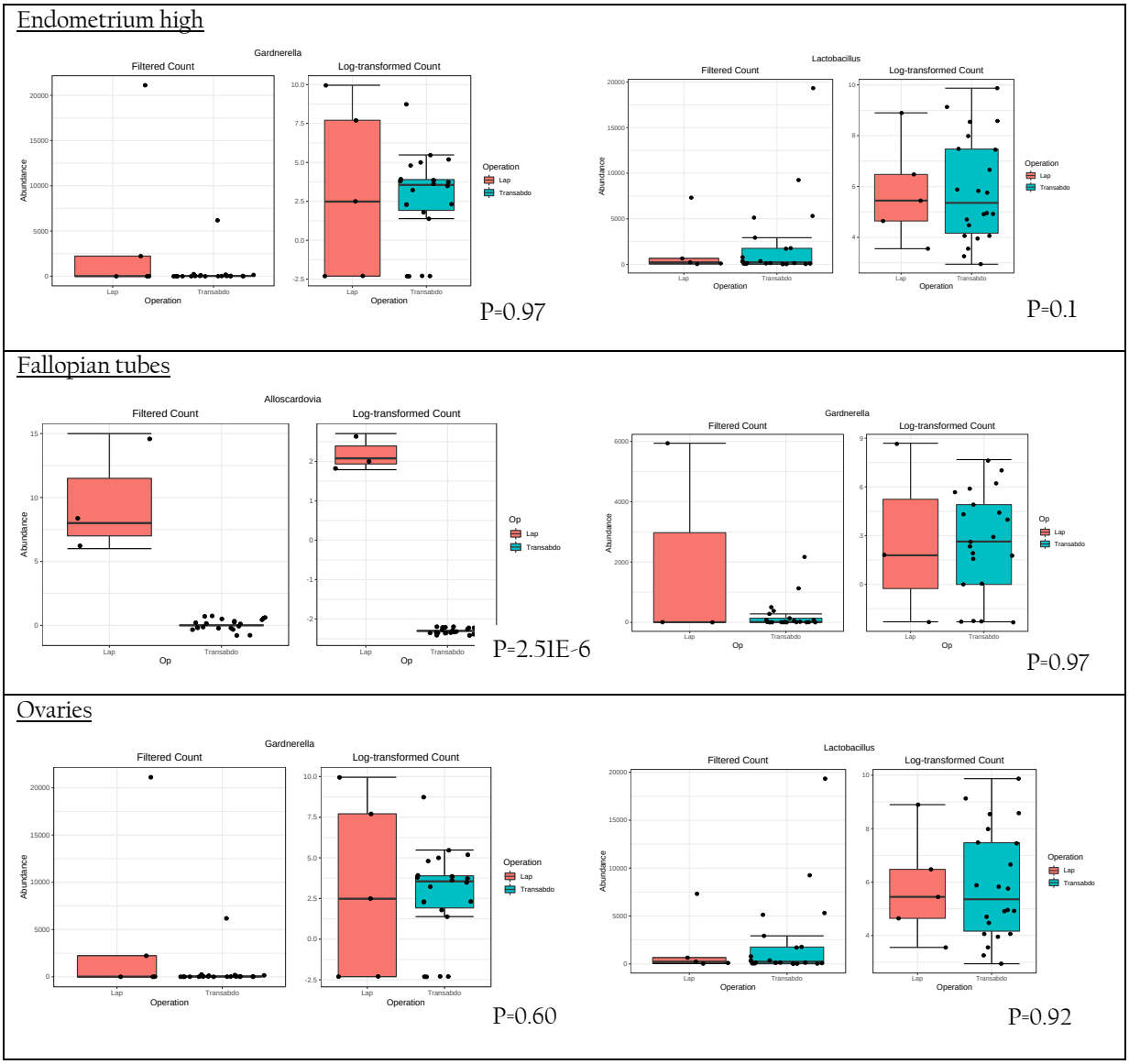


Figure 4.6 Univariate comparison analysis of laparoscopic and transabdominal samples by body site. Gardnerella and Lactobacillus were specifically reviewed as they are the most numerous microbes in the vagina and would be expected to be found in the upper genital tract should there have been significant contamination during a laparoscopic procedure. There was no statistical difference between the methods of operation bar the fallopian tubes, which showed a significant difference in the levels of Alloscardovia (p=2.51E-6).

4.3.3 The upper genital tract contains bacterial metataxonomic signatures above that of background controls

Comparison of relative bacterial species abundances between benign and control samples indicated that a bacterial signature could be detected above background levels in the upper genital tract sites when a threshold of P<0.05 was applied at species level (Mann-Whitney/Kruskall- Wallis, Table 4.3). Ninety percent of endometrial samples demonstrated a bacterial presence above background (median;

75% of samples using the mean). Thirteen species had a P value ‹0.05 however none withstood false discovery rate (FDR) ‹0.05 correction. In the fallopian tubes, 91% (median) and 82% (mean) of benign samples demonstrated a bacterial 16S rRNA signature above background. Of the 14 species with a P‹0.05, four also had an FDR ‹0.05. Ninety-four percent of ovarian benign specimens (median and mean) indicated a bacterial presence above background level. Twenty-one species had a P‹0.05 of which eight also had an FDR ‹0.05.

Twenty-six species were found to be significantly present in the ovarian malignant samples above background level (P‹0.05), with four also having an FDR ‹0.05 including *Azonexus_unclassified*, *Diaphorobacter_aerolatus*, *Novosphingobium_ aromaticovorans* and *Anaerococcus_unclassified*. Their clinical significance is currently undetermined.

Table 4.3 Species enriched in upper genital tract sites versus controls. The shaded results also met the FDR threshold (P ‹0.05).

NAME	P value*	NAME	P value*
Endometrium (benign vs controls)			
Lactobacillus crispatus	0.0036681	Catonella unclassified	0.042728
Anaerococcus unclassified	0.0070096	Lactobacillus leichmannii	0.042728
Corynebacterium ureicelerivorans	0.013101	Ezakiella unclassified	0.043653
Acinetobacter guandongensis	0.013114	Lactobacillus gasseri	0.043966
Alloscardovia omnicolens	0.013114	Gardnerella vaginalis	0.035579
Cornebacterium unclassified	0.013114	Nakamurella unclassified	0.042649
Haemophilus parainfluenzae	0.023894		
Fallopian tubes (benign vs controls)			
Acinetobacter_guangdongensis	4.2196E-5	Actinomyces_unclassified	0.014788
Novosphingobium_aromaticivorans	1.8983E-4	Finegoldia_magna	0.034989
Cloacibacterium_normanense	5.5085E-4	Comamonas_jiangduensis	0.034989
Azonexus_unclassified	0.0017625	Peptoniphilus_gorbachii	0.039661
Comamonas_denitrificans	0.002316	Campylobacter_ureolyticus	0.039661
Facklamia hominis	0.0052601	Alloscardovia_omnicolens	0.039661
Diaphorobacter_aerolatus	0.0064512	Lactobacillus_jensenii	0.049879
Ovary (benign vs controls)			
Acinetobacter_guangdongensis	2.5527E-4	Corynebacterium_tuberculostearicum	0.022101
Diaphorobacter_aerolatus	7.063E-4	Staphylococcus_caprae	0.023655
Lactobacillus_crispatus	0.0013262	Corynebacterium_appendicis	0.027041
Cloacibacterium_normanense	0.0014223	Papillibacter_unclassified	0.027041
Comamonas_denitrificans	0.0018051	Pseudorhodobacter_aquimaris	0.027041
Novosphingobium_aromaticivorans	0.0020981	Pseudoflavonifractor_unclassified	0.027076

Corynebacterium_ureicelerivorans	0.0031294	Prevotella_timonensis	0.028888
Azonexus_unclassified	0.0031342	Comamonas_jiangduensis	0.032732
Streptococcus_anginosus	0.004867	Finegoldia_magna	0.041779
Anaerococcus_unclassified	0.0066362	Ezakiella_unclassified	0.041793
Ralstonia_insidiosa	0.012197		
Ovary (malignant vs controls)			
Azonexus_unclassified	1.4969E-4	Comamonas_denitrificans	0.002977
Diaphorobacter_aerolatus	2.1367E-4	Acinetobacter_guangdongensis	0.004080
Novosphingobium_aromaticivorans	6.0988E-4	Cloacibacterium_normanense	0.004080
Anaerococcus_unclassified	0.0011883	Gardnerella_vaginalis	0.0042628

*Univariate analysis with Mann- Whitney/Kruskall Wallis test and p value cut off 0.05.

Further investigation assessing how consistent these results were in the whole cohort was performed (i.e., how commonly these species were seen in individual participants). The presence of the bacterial species with the most significant results in each site across all of the patient samples was examined. In the endometrium, three of the twenty benign endometrial samples did not display *Lactobacillus_crispatus* in levels above the contaminant control samples. The remaining 17 samples had fewer than 200 read counts. Twelve of the twenty samples did not have levels of *Anaerococcus_unclassified* above the control samples. Similar findings were noted in the fallopian tubes and ovarian samples, with at least two of the 11 benign fallopian tube samples, and occasionally as many as eight of the 16 benign ovarian samples not displaying read counts above 200 of the previously identified species reputed to be a 'true' biological sequence.

4.3.3.1 Sequence read counts decrease when ascending the genital tract

Using 16S rRNA sequencing as a proxy for bacterial load, sequence read counts were noted in each anatomical site in both benign and malignant samples.

Across both the benign and malignant samples, as expected, the sequence read counts were observed to be highest in the vagina (median sequence reads per sample 99,038). The read counts generally decreased as genital tract sampling ascended. Interestingly, the endometrial samples displayed numerically fewer sequence reads than the fallopian tubes and ovaries (median sequence reads per sample 1690, Table 4.4).

When analysing the samples by histology, overall, there was no difference in the sequence read counts. Whilst the malignant vaginal samples read counts were lower than the benign vaginal samples (median, t-test P= 0.0485), in the upper genital tract, despite numerical differences being noted between

histological types, these did not reach statistical significance (Kruskall- Wallis with Dunn's post hoc test P >0.9999, Table 4.4).

Table 4.4 16S rRNA sequence reads by anatomical body site and histology. Whilst not reaching statistical significance, overall, in the malignant samples, there are fewer read counts in the lower genital tract and more read counts in the upper genital trac. Specifically, the median of the malignant endometrial samples is the lowest read count of all the body sites.

SITE	Total read counts	Benign	Malignant	P value *KW, Dunns	Total No of samples
Vagina	2898907	1452855	1446052	0.0485	30
Min		19171	1866	(t test)	
Median		105035	88859		
Cervix	1659321	854159	805162	0.9672	29
Min		3549	1602		
Median		85977	33980		
Endometrium	282129	117087	165042	>0.9999	53
Min		159	217		
Median		2565	1528		
Fallopian tube	188909	43476	145433	>0.9999	24
Min		422	131		
Median		3996	5226		
Ovary	234038	112113	119414	>0.9999	40
Min		109	154		
Median		2371	2655		

4.3.4 Patients with malignant and benign conditions displayed common microbes throughout their genital tract.

The presence of a continuum of bacteria throughout the genital tract was next analysed. A continuum was defined as the presence of bacterial species in all sites of the LGT and UGT (vagina, cervix, endometrium, fallopian tubes and ovaries) at a relative abundance of at least 0.5%. Twelve of the 15 (80%) patients with malignant disease whose entire genital tract was sampled had at least one common bacterium present throughout their genital tract. Generally, the patients with malignant disease had a more diverse bacterial signature than the patients with benign disease. *Lactobacillus* was found throughout the genital tract in six of the 12 patients (50%) with malignant disease, whilst *Gardnerella* was found in three of the 12 (25%) patients (Figure 4.7). Other frequently seen genera include *Finegoldia*, *Prevotella*, *Staphylococcus*, *Corynebacterium*, *Bifidobacterium*, *Actinobacter*, *Bacteroides* and *Streptococcus*.

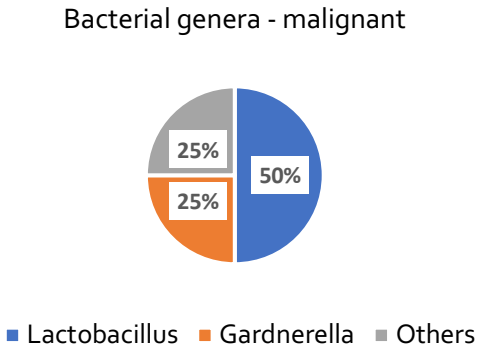


Figure 4.7 Percentage of patients from the cancer cohort who displayed a common bacterium throughout the genital tract. 50% had Lactobacillus, 25% had Gardnerella, with the remaining 25% displaying a range of other bacteria.

All eight of the benign patients displayed one or more common bacteria in their genital tract (Figure 4.8). Five of the eight women with benign disease (62.5%) had *Lactobaccilus* throughout their genital tract. Three of the eight women (37.5%) had *Gardnerella. Prevotella*, *Atopobium* and other genera were also noted.

In 87.5% (7/8) of benign patients and 66.7% (10/12) of malignant patients, the most abundant species in the LGT were also recovered from all sites in the UGT. This suggests that the easily accessible LGT could be used as a proxy to predict UGT microbial communities.

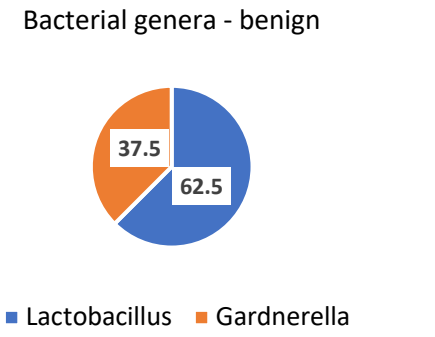


Figure 4.8 Percentage of patients from the benign cohort who displayed a common bacterium throughout the genital tract. 37.5% had Gardnerella whilst 62.5% had Lactobacillus at each body site.

All the malignant specimens with a bacterium in common throughout the genital tract were removed transabdominally (n/N 12/12), whilst three of the eight benign specimens with a common bacterium in all anatomical sites (n/N 3/8) were removed laparoscopically and five (n/N 5/8) transabdominally. There

was no discernible pattern to the amounts of bacterial species analysed, with some patients with a common bacterium throughout found in higher amounts in the upper genital tract locations, and others with higher amounts found in the lower genital tract.

The absolute read counts also varied widely. Whilst some sites with the common bacteria had read counts in the thousands, several sites reported single digit read counts. Representative benign and malignant cases are presented in Figure 4.9 a and b below, with the graphs for the remaining patients found in the appendices. The graphs illustrate the number of shared species that overlap in the reproductive tract, the percentage relative abundance and the read counts of the shared species in each anatomical site.

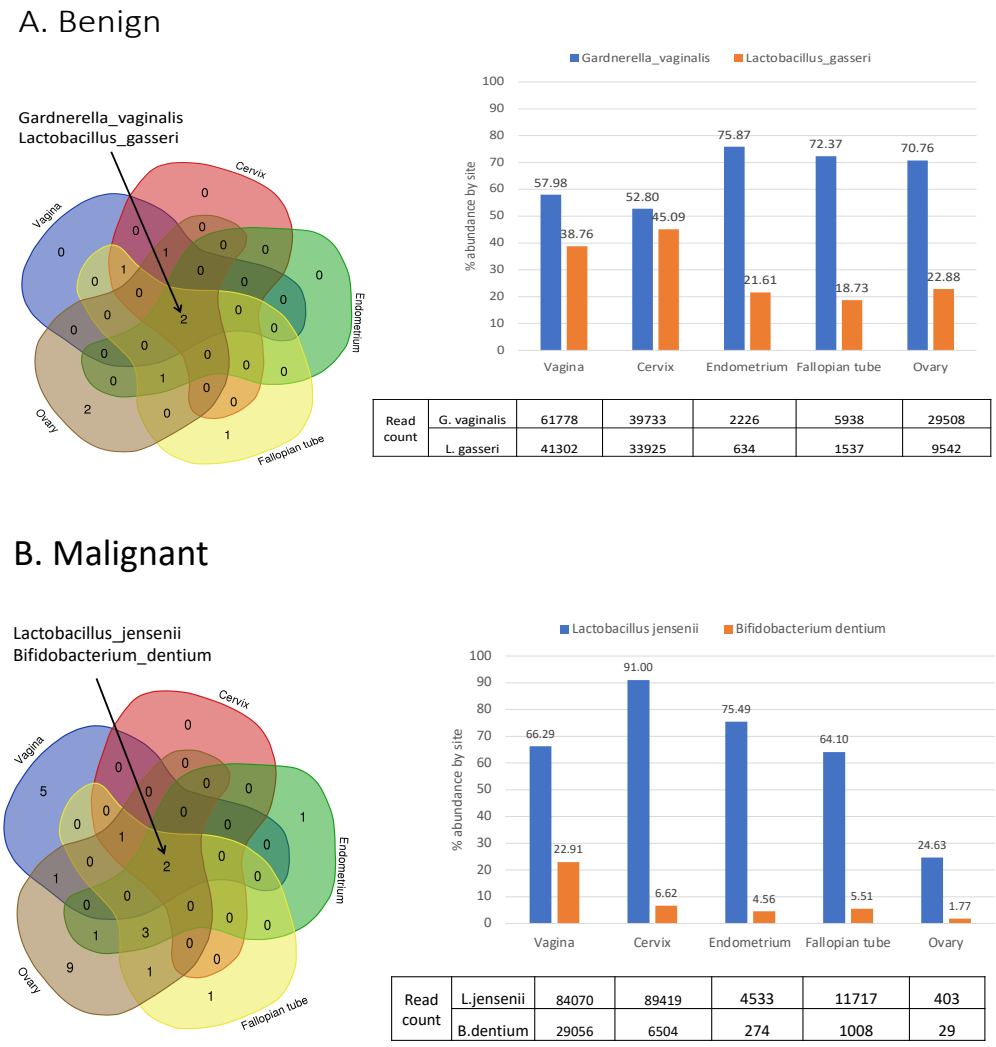


Figure 4.9 Venn diagrams illustrating the number of shared species throughout the genital tract of a benign (A; EW114) and malignant (B; EW19) patient. The percentage relative abundance and read counts of the common species are also demonstrated.

Exploring the theory of a common species or continuum throughout the genital tract further, it was found that although the common species sometimes represented a large proportion of the total amount of bacterial signature in a genital tract (see Figure 4.9a and b), they were often a fraction of the total number of species within the genital tract, particularly in patients with malignant disease, indicating high diversity (Table 4.5).

Table 4.5 The relative abundance of the shared species found throughout the reproductive tract as a percentage of the total number of unique species per patient. No significant difference was noted between the benign and malignant samples.

Benign	n/N, %	Mean, median (%)	Malignant	n/N, %	Mean, median (%)
EW90	1/57, 1.75	7.4, 4.6	EW6	1/33, 3	3.5, 2.9
EW71	1/ 17, 5.88		EW99	2/34, 5.88	
EW41	1/9, 11.11		EW105	3/49, 6.12	
EW33	1/16, 6.25		EW107	2/71, 2.82	
EW82	1/57, 1.75		EW128	1/69, 1.45	
EW124	1/51, 1.96		EW78	1/53, 1.89	*Three patients with malignant disease did not display a common species throughout the genital tract
EW114	2/8, 25		EW19	2/25, 8	
EW103	2/59, 3.39		EW66	1/34, 2.94	
			EW18	1/83, 1.2	
			EW73	2/52, 3.85	
			EW62	1/46, 2.17	
			EW91	1/56, 1.79	

4.3.5 The endometrium and UGT shared a greater proportion of bacterial species than the LGT and the endometrium.

Continuing to examine the concept of a common species but taking a broader view, the presence and subsequent degree of overlap of bacterial species between each body site for the entire cohort, and subsequently the three 'compartments' of the genital tract; the lower genital tract comprising the vagina and cervix, the intrauterine space i.e., endometrium, and the upper genital tract comprising the fallopian tubes and ovaries was explored.

There was a higher degree of species commonality between the endometrium and UGT (68.8%) than the LGT and the endometrium (42.6%) (Figure 4.10, 4.11) when comparing all three compartments. This difference was also noted when the compartments were compared directly (LGT vs endometrium

44.44%, endometrium vs UGT 72.39%, Figure 4.10, Table 4.6). The fallopian tube and ovaries were also noted to have more species in common that the vagina and cervix (Figure 4.11, Table 4.6).

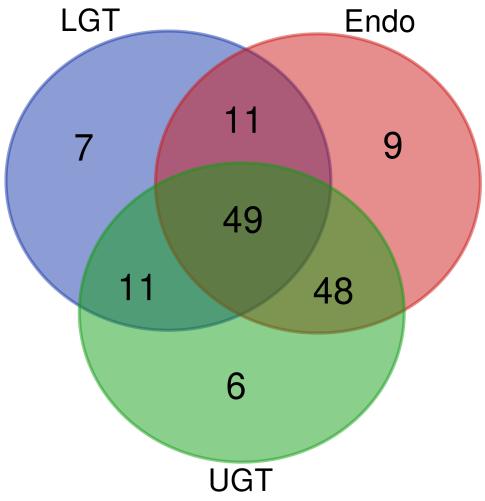


Figure 4.10 Venn diagram of the number of bacterial species identified which are shared between the lower genital tract, endometrium and upper genital tract. The endometrium and UGT share a higher number of bacterial species than the LGT and endometrium (68.8% vs 42.6%). The overall percentage relative commonality between the three compartments is 34.75%.

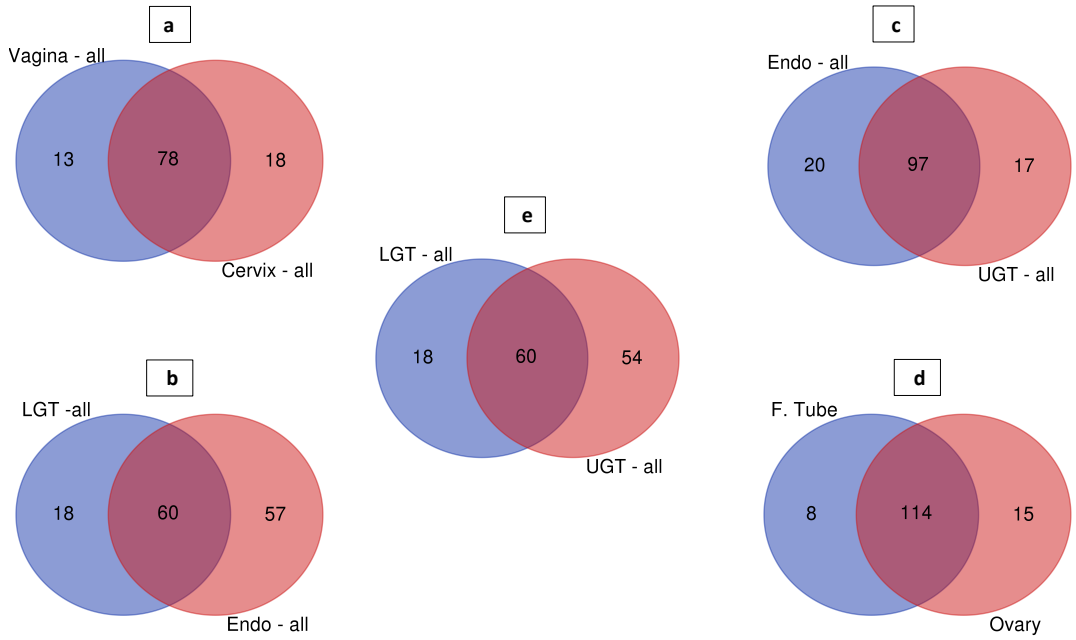


Figure 4.11 Number of species that overlap at different anatomical sites and compartments. All samples per site included for analysis. a. Vagina vs cervix. b. Lower genital tract (vagina and cervix) vs endometrium. c. Endometrium vs upper genital tract (fallopian tube and ovary) d. Fallopian tube vs ovary. e. Lower genital tract vs upper genital tract.

Table 4.6 Percentage relative abundance of common species between anatomical sites when compared directly. The fallopian tube and ovary have a higher percentage of commonality than the vagina and cervix. The endometrium and UGT also have a higher degree of commonality than the LGT and endometrium.

Anatomical site/compartment (all samples)	Relative abundance of shared species (%)
a) Vagina vs Cervix	71.56%
b) LGT vs Endometrium	44.44%
c) Endometrium vs UGT	72.39%
d) Fallopian tube vs Ovary	83.21%
e) LGT vs UGT	45.45%

4.4 DISCUSSION

To date, literature describing the 'microbiome' of the female genital tract has largely focused on the cervicovaginal niche both due to its easy accessibility and its high volume of biomass. Previous methods of assessing the microbial landscape relied largely upon classical methods including microscopy and culturing of bacteria, at least of those species which were amenable to culture. Since the advent of molecular based approaches like NGS, however, there is greater appreciation of bacterial diversity in different biological niches. Yet the sensitively of these means that their application to low biomass samples have been impeded by concerns regarding contaminants leading to overestimation of the true microbial content[270]. This was further compounded by the biologically implausibility of finding specific bacteria in certain body niches, e.g., species that require photosynthesis identified in the placenta [270].

Many contaminants derived from environmental (e.g., laboratory, reagents) and patient-specific (e.g., blood, skin) sources have since been identified with calls for more stringent implementation of contaminant control to help minimise this risk of background 'noise' which may dilute the true biological data. Maintaining absolute sterility is difficult due to the nature of sample collection and processing thus to address these issues of contamination, I implemented several different control samples to allow for assessment and removal of potential contaminants from environmental e.g., airborne, and material e.g., kit, equipment, sources during the sample collection and processing. The value of these controls was evident as their inclusion resulted in 20.45% of the total sequencing data being highlighted as potential contamination.

Participants can be sources of cross-contamination as evidenced by the six samples removed due to intraoperative contamination secondary to bowel resection during the collection process. Another potential source of cross-contamination specific to this cohort relates to intraoperative practices of laparoscopic hysterectomies, where an intrauterine manipulator is inserted through the vagina and

cervix into the uterine cavity, and the surgical specimen is removed through the vaginal incision. It was important to address this considerable contamination risk and intra-patient correlation of vaginal and endometrial samples was performed on the patients who had laparoscopic procedures, with further comparison of laparoscopic and open procedures.

Comparison of the five laparoscopic and 24 transabdominal hysterectomies did not reveal a significant translocation of vaginal microbes to the upper genital tract. This is in line with a previous study which compared samples collected via transcervical and transuterine methods and demonstrated that they were highly correlated[106]. The imbalance of procedure type reflects the predominantly transabdominal nature of ovarian cancer surgery.

With the 'true' sequencing data, I next tested the hypothesis of whether a true bacterial signature above background noise was present in the UGT including the endometrium, fallopian tubes and ovaries. Compositional analysis demonstrated that all >90% of samples from all three 'compartments' of the UGT were enriched compared to controls. Ascending from the benign endometrial samples, species including *Lactobacillus_crispatus*, *Anaerococcus_unclassified*, *Lactobacillus_leichmannii*, *Gardnerella_vaginalis* were identified; *Acinetobacter_guandongensis*, *Novosphingobium_aromaticivorans*, *Cloacibacterium_normanense*, *Facklamia_hominis* were identified in the benign fallopian tube samples whilst *Lactobacillus crispatus*, *Acinetobacter_guangdongensis*, *Comamonas_dentrificans* were identified in the benign ovarian samples amongst several others. Chen also noted *Facklamia* in the fallopian tubes of women with benign uterine disease[106]. To date no studies have noted the presence of *Acinetobacter_guandongensis* in the FGT however *Acinetobacter lwoffi* has been found in malignant ovarian tissues by Zhou and Qin[34,38]. Another member of the Acinetobacter genus, *Acinetobacter baumanii*, has been reported to secrete lipopolysaccharides and extracellular vesicles that activated the Toll-like receptor 4 (TLR-4) signaling cascade and produces interleukin 8 and tumour necrosis factor alpha (TNFα) release from THP-1 cells (in vitro cancer cell models)[271].

This study further reinforced current distrust regarding identification of a genuine biological signature in these low biomass niches. Interrogation of the compositional data comparing the benign samples and control samples revealed that 15-56% of benign samples in the UGT did not have measurable amounts of the putative enriched species. Several known contaminants were also identified in our study and other studies despite the efforts made to minimise the effect of external contaminants.

A similar trend as to the current published literature was noted in our study regarding the volume of sequencing data. Overall, the number of read counts decreased as the genital tract was ascended from the vagina to the ovaries. Several specific results should be noted however. The vaginal samples from

women with ovarian cancer had fewer read counts than those with benign conditions (median). Whilst this did not reach statistical significance (P=0.049), it was a trend that was seen again in the cervix and further upstream in the endometrium of women with ovarian cancer (median). Interestingly, this trend reversed in the UGT where the read count was higher in the fallopian tubes and ovaries of women with ovarian cancer. One theory could suggest that the presence of bacteria in the fallopian tubes of women with ovarian cancer might indicate that they instigate the chain of events involving the process of inflammation, development of STIC lesions and subsequent fallopian tube carcinoma[82,84,112,272]. Alternatively, their presence could imply that they are a (by)product of oncogenesis or that they are transient colonisers. Lastly, and the most plausible explanation, is that the result is spurious and due to contamination that was not addressed by our methods.

Once we had established that a genetic signature was present throughout the genital tract, I next focused on whether a solid continuum of bacterial sequencing signal was present at each body site per patient and if there was a correlation between the LGT and UGT. Eighty percent (n/N, 12/15) of the patients with ovarian cancer were found to have at least one species present in all body sites. At genus level, Lactobacillus was present in 50% of those women, and Gardnerella in 25%, amongst several other genera. All (n/N, 8/8) of the women with benign conditions demonstrated the presence of at least one species at each body site. 62.5% of the women had Lactobacillus, whilst 37.5% had Gardnerella.

Whilst these results appear to be noteworthy, upon further analysis, huge variation in the read counts at each body site was seen. In most patients, the expected distribution of read counts was observed i.e., large volume of read counts in the vagina decreasing in volume as the genital tract is ascended as noted in several studies. Only five of the 23 patients who had their entire genital tract sampled had sequence read counts >200 at each body site for the 'common' microbe. Many had double or single digit counts in the UGT sites.

I also reviewed the relative abundance of the 'common' species, and whilst it often represented the highest proportion of the sequencing data, it was only one of many species that were detected. These results may indicate a host specific microbe effect as reported by Walther- Antonio et al, or 'microbial migration' from the LGT as noted by Qin et al, or via a different route e.g., haematogenous[38,105,273,274].

Finally, the degree of overlap between each 'compartment' was examined i.e., the LGT, the endometrium and the UGT. A higher degree of overlap of species was noted between the endometrium and the UGT than the LGT and the endometrium when all three compartments were examined simultaneously. This was reconfirmed when the compartments were examined directly. The highest degree of overlap between sites was noted between the fallopian tube and ovaries, although unsurprisingly the vagina and cervix also displayed a high degree of similarity. Whilst the concept of

the cervical mucous barrier as a completely impenetrable blockade has been discounted, it still plays a valuable role in limiting the translocation of microbes from the LGT to the UGT by producing a highly hydrated gel that protects against microbial attacks along with secreting many defensive substances like defensins, mucins and lyzozyme[275]. Another theory which may explain why a higher degree of correlation was noted between the endometrium and UGT rather than the LGT and endometrium may be related to the consistent shedding of the endometrial lining throughout the female reproductive lifespan; microbes may ascend to the UGT but due to the constant changeover and 'flushing' out of the endometrium, have less interference than the LGT.

4.5 STRENGTHS AND LIMITATIONS

In this study, a comprehensive multifaceted approach was taken to minimise the risk of contamination from non-biological sources, and to decontaminate the resulting data using a robust standardised tool. Previous studies of low biomass sites including those on gynaecological malignancies, have been hampered by concerns regarding contamination of data resulting in misleading conclusions. This approach highlighted the high volume of contaminant data and underlined the importance of having strict procedures in place to limit contamination 'noise' when undertaking studies into low biomass sites. In this study, read counts were used as a proxy for bacterial load which gives a broad view of biological mass, however, is not fully representative as whilst sequence counts indicate the presence of a bacterial species, it does not account for the number of copies of that species. As such, these results should be viewed with caution.

4.6 CONCLUSIONS

This study highlighted how imperative it is to undertake systematic comprehensive steps to ensure that non-source contaminants are minimised and resulting data is optimally de-noised. A considerable proportion of the initial results, particularly in the UGT were found to be contaminants which could have led to spurious results if not identified and removed.

Correlation analysis indicated that minimal translocation of biological material occurred when specimens were removed laparoscopically versus transabdominally. A bacterial signature was noted both compositionally and quantitively in the UGT above the background contaminants. Migration analysis demonstrated that 80% of cancer samples and 100% of benign controls had a bacterial continuum with at least one species noted throughout the genital tract with 75% of cancer and 100% of benign samples displaying vaginal commensals throughout the genital tract. Analysis of the three

'compartments' demonstrated that the LGT and endometrium shared fewer bacterial taxa than the endometrium and the UGT. The fallopian tubes and ovaries were enriched with the highest number of shared taxa, with the LGT and endometrium the fewest.

4.7 STATEMENT OF CONTRIBUTION

Dr Sherrianne Ng completed the processing of sequencing data for identification of contaminant sequences using the Decontam package.

CHAPTER FIVE: COMPARISON OF FEMALE GENITAL TRACT BACTERIAL PROFILES IN PATIENTS WITH OVARIAN CANCER AND BENIGN CONTROLS

5.1 INTRODUCTION

Since Pasteur and Koch first recognised the pathogenic effect of microbes, attempts have been made to elucidate the interplay of microbes in human health and disease. Several conditions are correlated with both microbes and alterations in the microbiome including inflammatory bowel disease[276], cancer[112] and obesity[277].

Relevant to the work described in this thesis, a landmark study by Ravel et al., in 2011 provided the first description of the asymptomatic vaginal microbiome[100,278]. Five different community state types (CSTs) were described of which four were *Lactobacillus* species (spp.) -dominant, with a fifth that was *Lactobacillus* spp.-deplete and highly diverse. Since then, a relationship between these *Lactobacillus* spp.-deplete type CSTs and reproductive tract infections (e.g., BV (bacterial vaginosis), urinary tract infections (UTIs)) and reproductive health (e.g., preterm birth risk, IVF success) has been firmly established [139,279-281]. BV, a dysbiotic state characterised by depletion of *Lactobacillus* spp. and increased diversity and enrichment of anaerobes, is correlated with human papilloma virus (HPV), which is well established as the cause of >99% of cervical cancers. Higher vaginal microbial diversity is associated with persistence of HPV infection and disease progression[104,282]. BV is also associated with an activated immune response in the vaginal mucosa. Most studies have reported that proinflammatory cytokines (e.g., IL1β), are elevated in women with BV, whereas antimicrobial peptides such as secretory leukocyte protease inhibitor (SLPI) are lowered. This proinflammatory state is also correlated with disruption of mucosal integrity due to protease secretion and epithelial cell wall injury which can increase potential for HPV infection[104]. Consistent with these findings, cross-sectional microbiota data has associated *Lactobacillus* spp. depletion and increased relative abundance of taxa including *Sneathia*, *Gardnerella*, *Atopobium*, *Bacteroides* and *Prevotella* in women with HPV infection, cervical intraepithelial neoplasia (CIN) and invasive cervical carcinoma[104].

Several recent studies have investigated the association of a putative upper genital tract microbiome with upper genital tract cancer risk and progression but these have proven challenging due to difficulties in accessing the upper genital tract and contamination issues relating to low biomass sites. Studies focusing on the endometrial cavity in both benign and malignant conditions[105,106,109,268,269], have mainly noted detection of *Lactobacillus* species in the benign endometrium, however in far lower amounts that in the cervicovaginal niches. In endometrial cancer samples, higher diversity and enrichment in genera including Flavobacterium, Cornebacterium, Bifidobacterium, Staphylococcus and Streptococcus and Bacteroides, Atopobium, Porphyromas, Shigella, Faecalibacterium and Micrococcus have all been reported[105,283].

Relatively fewer studies have attempted to identify microbes in the fallopian tubes or ovaries. Initially these studies focused on investigating the possible presence of microbes in the upper genital tract of healthy women[106,109,143,284]. More recently, studies investigating specific microbes related to ovarian cancer[28,29] and the presence of a potential microbiome or microbial colonisation in ovarian cancer[32-35] have been reported. Finally, attempts to explore the mechanisms by which microbes or microbial patterns might interact with the metabonome to potentiate carcinogenesis[34] were described. Both the fallopian tubes and ovaries have displayed high diversity with Firmicutes, Actinobacter and Proteobacter all featuring.

5.2 AIMS

The aims of this chapter are to:

1) Compare the bacterial composition of benign and malignant samples throughout the female genital tract
2) Determine if ovarian cancer is associated with a specific cervicovaginal bacterial signature

5.3 RESULTS

5.3.1 Lactobacillus species depletion of the female genital tract is associated with ovarian cancer.

Hierarchical clustering analysis did not determine any statistically significant patterns at genera nor species taxonomic level according to relative abundance when attempting to differentiate the benign and malignant samples at all anatomical sites throughout the female genital tract. The samples were reviewed according to anatomical site, within which subcategories including menopausal status, BMI, age, ethnicity, type of operation, PID, HRT or OCP use were also analysed. No statistically different results were noted relative to the cohort metadata so were not included in the results.

Low read counts due to rarefying was particularly an issue in the UGT sites, with some of the clustering being performed on samples which had been rarefied to 100 read counts e.g., ovarian samples.

As previously mentioned, two methods of clustering analysis were implemented. Both Ward linkage and Euclidean distance and VALENCIA were used for the vaginal and cervical samples but the UGT

samples were analysed using Ward linkage and Euclidean distance only. The VALENCIA method is only relevant to LGT samples as the tool was developed based on a cohort of vaginal samples only.

5.3.1.1 Vagina

Examination of the vaginal bacterial profiles at genus level (Figure 5.1, Table 5.1), indicated that the benign samples could be clustered into three main groups characterized by relative abundance of Lactobacillus or Gardnerella; Lactobacillus dominant (4/10, 40%), Gardnerella/ Lactobacillus dominant (3/10, 30%) and high diversity (3/10, 30%). A trend increase in prevalence of Lactobacillus-depleted communities was observed in malignant samples compared to benign samples but this did not reach significance (11/18, 61.1%; Chi-squared P=0.2432).The high diversity group, into which most of the malignant samples were clustered, was characterized by Lactobacillus depletion and included samples with higher Shannon α- diversity (Figure 5.13) than the other two clusters and also a small number of samples dominated by one or two bacterial species not seen in other samples.

When sequencing data was clustered at species level using the VALENCIA nearest-centroid algorithm, 70% (7/10) of benign samples were clustered to a Lactobacillus-dominant CST (CST I-III, V), whereas 61.1% of malignant samples clustered to Lactobacillus-deplete CST IV (Fishers exact P=0.2365).

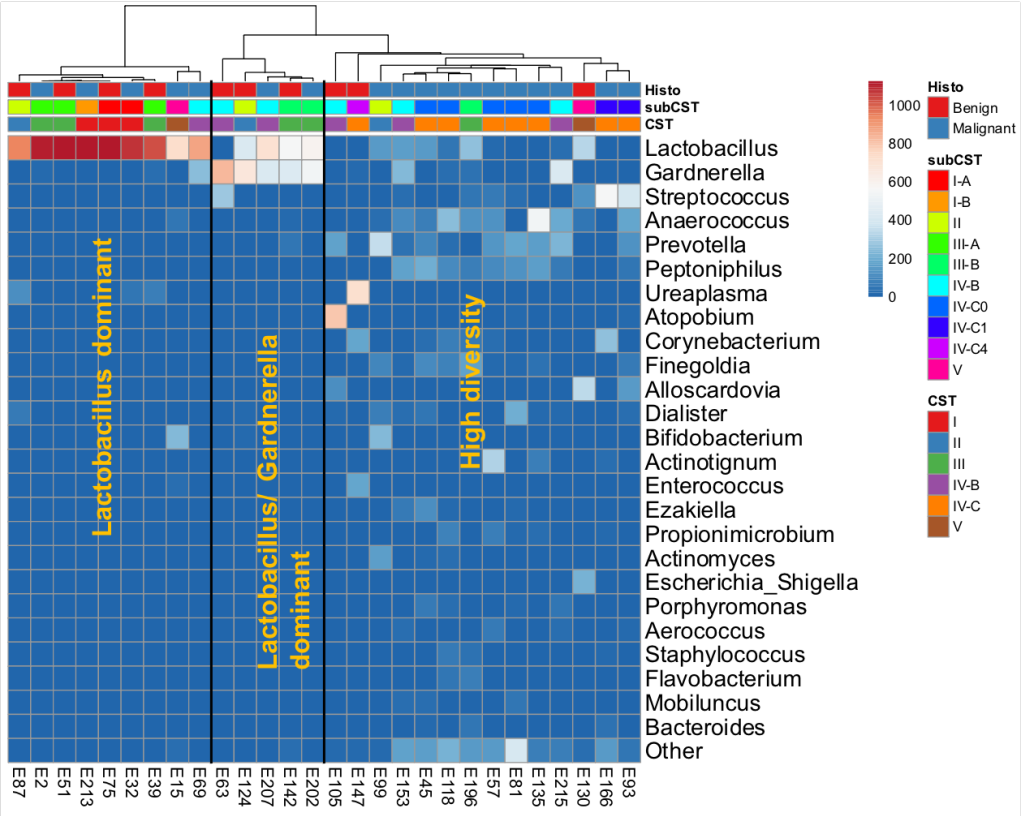


Figure 5.1 Hierarchal clustering analysis of all vaginal samples using both Ward linkage and Euclidean distance and the VALENCIA method at genus level.

Table 5.1 Results of hierarchal clustering of vaginal samples using a) Euclidean distance and Ward linkage and then using b) VALENCIA method

A. Vagina	Lactobacillus dominant n/N (%)	Lactobacillus/ Gardnerella dominant n/N (%)	High diversity n/N (%)	Total n/N (%)
Benign	4/10 (40)	3/10 (30)	3/10 (30)	10/10 (100)
Cancer	5/18 (27.8)	2/18 (11.1)	11/18 (61.1)	18/18 (100)
Total	9/28 (32.1)	5/28 (17.9)	14/28 (50)	28/28 (100)
*Chi-squared (2.83)				P=0.2432

B. Vagina	CST I-III, V (Lactobacillus dominant) N/n (%)	CST IV (Lactobacillus deplete) n/N (%)	Total n/N (%)
Benign	7/10 (70)	3/10 (30)	10/10 (100)
Cancer	7/18 (38.9)	11/18 (61.1)	18/18 (100)
Total	14/28 (50)	14/28 (50)	28/28 (100)
*Fishers exact			P=0.2365

At species level (Figure S2 in appendices), four groups are seen using Ward linkage; *L. crispatus* dominant (1/10, 10%), *L. iners* dominant (2/10, 20%), Lactobacillus- deplete (4/10, 40%) and *L. iners/G.vaginalis/L. fornicalis* dominant (3/10, 30%). Most of the malignant samples are Lactobacillus deplete (12/18, 66.7%), with 11.1% (2/11) clustering to *L. crispatus*, 1/18 (5.6%) to *L. iners* and 3/18 (16.7%) to *L. iners/G. vaginalis/L. fornicalis*. These results are non-statistically significant (P= 0.4588). The VALENCIA results reflect the previous results with 70% (7/10) of benign samples and 38.9% (7/18) of malignant samples clustering to Lactobacillus- dominant CSTs, and 30% (3/10) of benign samples and 61.1% (11/18) of malignant samples clustering to the Lactobacillus-deplete CST (Fishers exact P=0.2365).

5.3.1.2 Cervix

At genus level, Ward clustering divided the cervical samples into four groups; Lactobacillus-dominant, Gardnerella-dominant, Lactobacillus/Gardnerella- dominant and Lactobacillus- deplete (Figure 5.2). The benign samples were subjectively Lactobacillus-dominant (4/10, 40%), with three out of ten samples (30%) in the Lactobacillus-deplete group (Chi-squared P=0.754) (Table 5.2). Applying the VALENCIA method, eight out of ten (80%) benign samples were grouped into Lactobacillus-dominant CSTs, with seven of 18 malignant samples in the Lactobacillus-dominant group, and 11 of 18 (61.1%) in the Lactobacillus-deplete group (Fishers exact test P=0.0546).

At species level, the benign samples reflect the genera distribution whilst the 9 of the 18 (50%) malignant samples are clustered into the Lactobacillus- deplete group (Chi-squared P=0.5928). The VALENCIA clustering also reflects the genera groups, with 8 out of ten (80%) benign samples in Lactobacillus-dominant CSTs, and 11 out of 18 (61.1%) malignant samples in the Lactobacillus-deplete CST (Fishers exact P=0.0545) (Figure S2 in appendices).

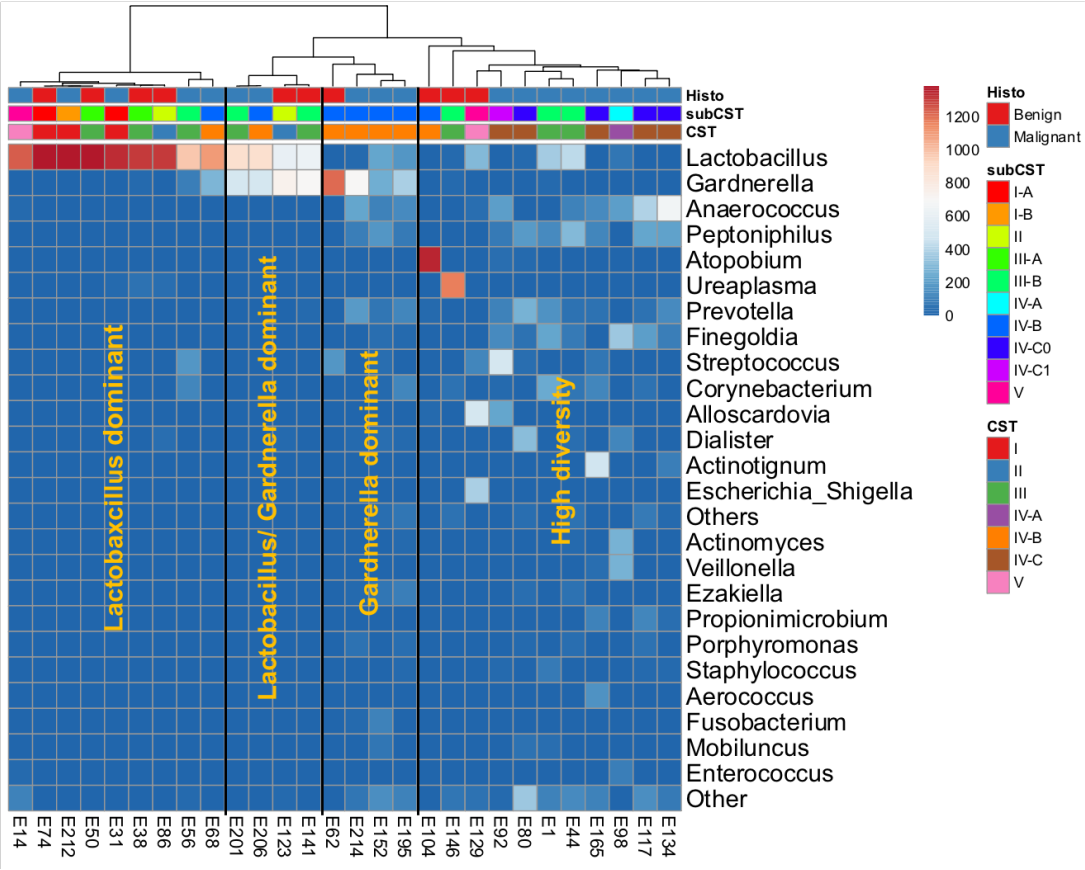


Figure 5.2 Hierarchal clustering analysis of all cervical samples using both Ward linkage and Euclidean distance and the VALENCIA method at genus level.

Table 5.2 Results of hierarchal clustering of cervical samples using a. Euclidean distance and Ward linkage and then b. VALENCIA method.

A. Cervix	Lactobacillus n/N (%)	Lactobacillus/ Gardnerella dominant n/N (%)	Gardnerella dominant n/N (%)	High diversity n/N(%)	Total n/N (%)
Benign	4/10 (40)	2/10 (20)	1/10 (10)	3/10 (30)	10/10 (100)
Cancer	5/18 (27.8)	2/18 (11.1)	3/18 (16.7)	8/18 (44.4)	18/18 (100)
Total	9/28 (32.1)	4/28 (14.3)	4/28 (14.3)	11/28 (39.3)	28/28 (100)
*Chi-squared =1.2					p=0.754027

B. Cervix	CST I-III, V (Lactobacillus dominant) N/n (%)	CST IV (Lactobacillus deplete) n/N (%)	Total n/N (%)
Benign	8/10 (80)	2/10 (20)	10/10 (100)
Cancer	7/18 (38.9)	11/18 (61.1)	18/18 (100)
Total	15/28 (50)	13/28 (50)	28/28 (100)
*Fishers exact			P=0.0.0546

5.3.1.3 Endocervical junction

Patterns of relative abundance initially appeared to reverse as the genital tract was ascended. At the endocervical junction, Ward-based hierarchal clustering subjectively grouped most of both the benign and malignant samples in the Lactobacillus-deplete group; benign, nine from 14 samples (64.3%), malignant, eight from 11 samples (72.7%) (Figure 5.3). These samples, taken from the lower aspect of the endometrium, did not show any statistically significant separation based on relative abundance of Lactobacillus, Gardnerella or any other genera. Five benign samples displayed subjectively higher levels of Lactobacillus (5/14, 28%) whilst six samples were solely Lactobacillus-deplete (6/14, 43%)(Chi-squared P=0.889) (Table 5.3). Three remaining samples clustered with the Lactobacillus-deplete group, but the read counts could similarly be said to subjectively reflect a Gardnerella- dominant group (3/14, 21%). Most of the malignant samples clustered in the Lactobacillus-deplete group (8/11, 73%). The group was titled Lactobacillus- deplete rather than high diversity to reflect that no statistically significant difference in intra (α) or inter (β)-sample diversity was noted (Figure S3 in appendices).

At species level, two benign samples and one malignant sample had read counts corresponding to dominant levels of *L. iners* with the remaining 11 samples generally high diversity and Lactobacillus-deplete (Chi squared= P=0.673) (Figure S2 in appendices).

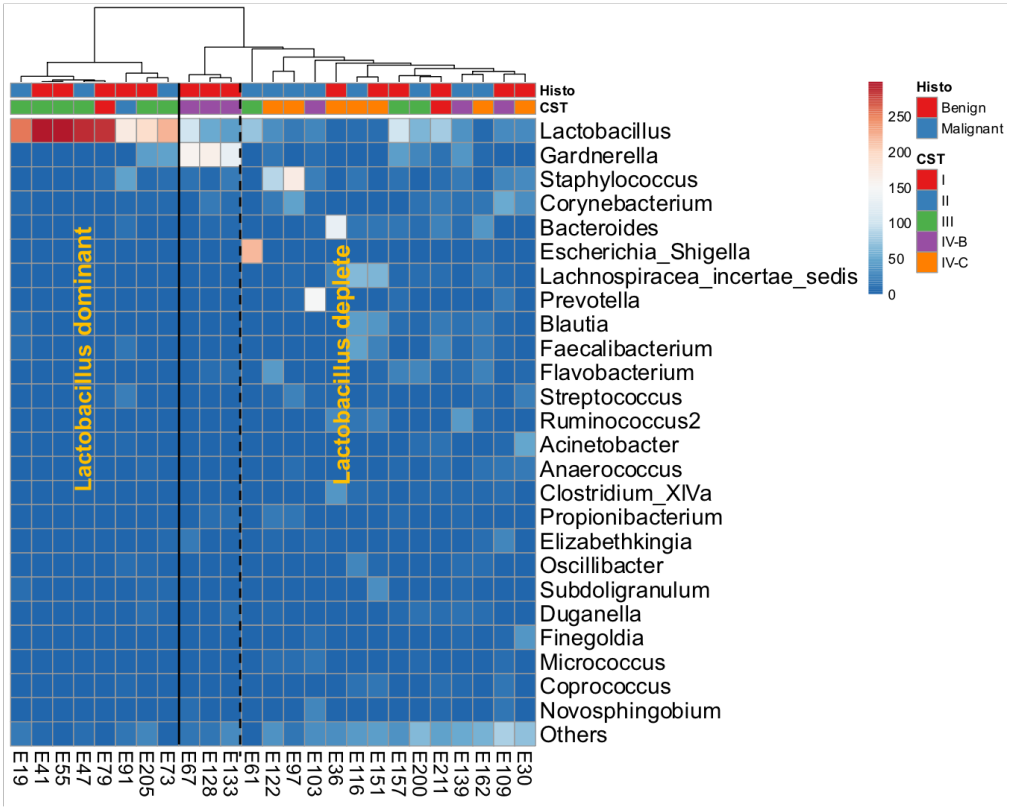


Figure 5.3 Hierarchal clustering analysis of all endocervical samples using Ward linkage and Euclidean distance and the VALENCIA method at genus level.

Table 5.3 Results of hierarchal clustering of endocervical samples using Euclidean distance and Ward linkage.

Endocervical	Lactobacillus-dominant n/N (%)	Lactobacillus- deplete n/N (%)	Total n/N (%)
Benign	5/14 (35.7)	9/14 (64.3)	14/14 (100)
Cancer	3/11 (27.3)	8/11 (72.7)	11/11 (100)
Total	8/25 (32)	17/25 (68)	25/25 (100)
*Chi-squared =0.7			P=0.402784

5.3.1.4 Fundus of endometrium

The pattern at the fundus of the endometrium at genus level changed again with only 36% of benign samples and 27% of malignant samples clustering into the Lactobacillus-deplete group (Figure 5.4, Table 5.4).

Three out of 11 benign samples clustered into a Lactobacillus-dominant group (27%), whilst four from 11 (36%) samples clustered into a Gardnerella-dominant group.

Interestingly, eight of the 11 (73%) malignant samples were grouped into a Lactobacillus-dominant group, with only three of the 11 (27%) samples in the Lactobacillus-deplete group. No malignant samples clustered into the Gardnerella-dominant group which prevented a P-value from being calculated (Table 5.4).

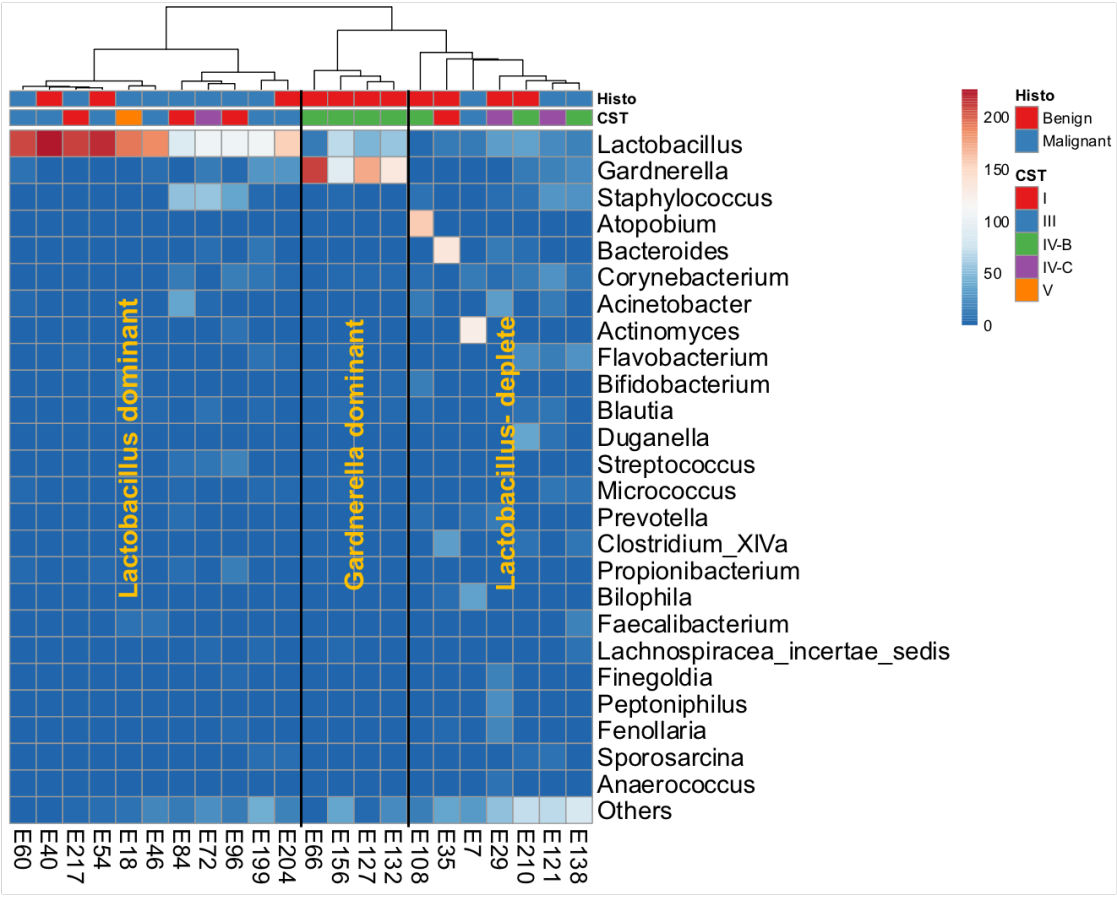


Figure 5.4 Hierarchal clustering analysis of all fundal endometrial samples using Ward linkage and Euclidean distance .

Table 5.4 Results of hierarchal clustering of fundal endometrial samples using Euclidean distance and Ward linkage

Fundus of endometrium	Lactobacillus dominant n/N (%)	Gardnerella dominant n/N (%)	Lactobacillus deplete n/N (%)	Total n/N (%)
Benign	3/11 (27.3)	4/11 (36.4)	4/11 (36.4)	11/11 (100)
Cancer	8/11 (72.7)	0/11 (0)	3/11 (27.3)	11/11 (100)
Total	11/22 (50)	4/22 (18.2)	7/22 (31.8)	22/22 (100)
*Chi-squared = unable due to 0.				

5.3.1.5 Fallopian tubes

Eight of the 11 (73%) malignant samples were grouped into the Lactobacillus-deplete group in when analysed. Five of the 11 benign samples also clustered into the Lactobacillus-deplete group, with two (18%) samples dominated by Gardnerella and four samples (35%) in the Lactobacillus-dominant group (Figure 5.5, Table 5.5).

At species level, three benign samples (27%) were *L.iners* dominant, with two samples *G. vaginalis* dominant. The remaining six benign samples (55%) and ten of the 12 malignant samples (83%) were Lactobacillus-deplete (Figure S2 in appendices).

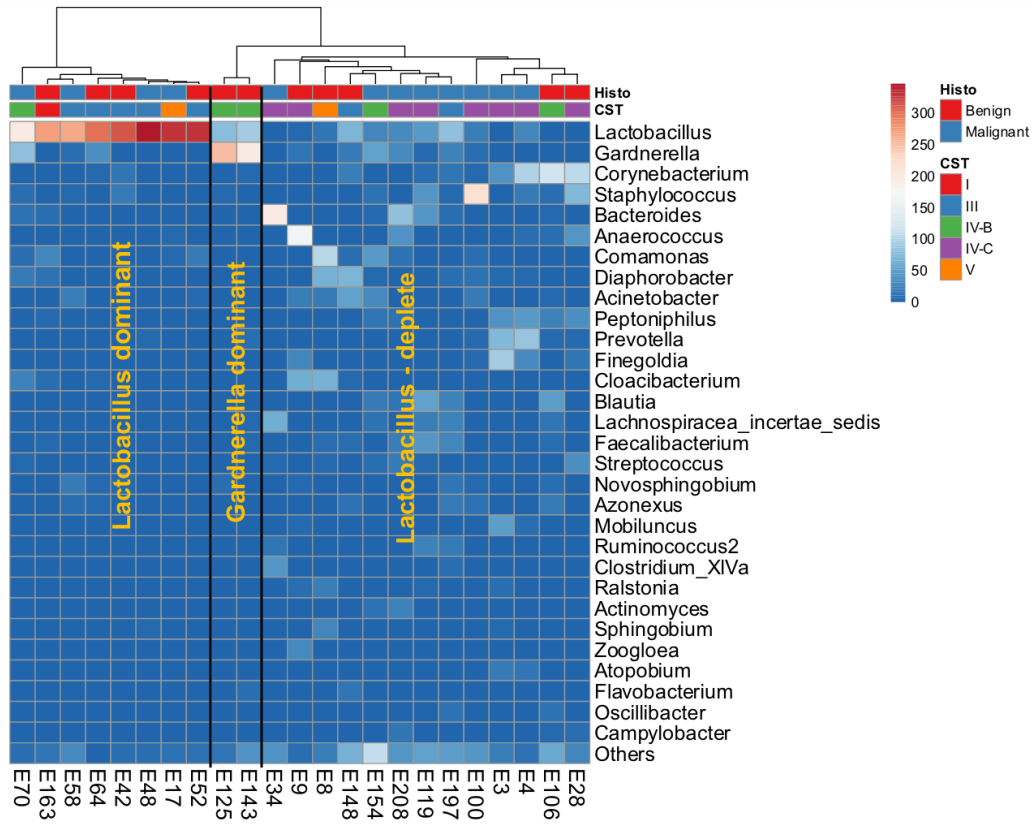


Figure 5.5 Hierarchal clustering of fallopian tube samples using Ward linkage and Euclidean distance.

Table 5.5 Results of hierarchal clustering analysis of fallopian tube samples using Euclidean distance and Ward linkage.

Fallopian tube	Lactobacillus dominant n/N (%)	Lactobacillus/ Gardnerella dominant n/N (%)	Lactobacillus deplete n/N (%)	Total n/N (%)
Benign	4/11 (35.4)	2/11 (18.2)	5/11 (45.5)	11/11 (100)
Cancer	4/11 (36.4)	0/11 (0)	8/11 (72.7)	11/11 (100)
Total	8/22 (36.4)	2/22 (9.1)	13/22 (59.1)	22/22 (100)
*Chi-squared = unable due to 0				P=.

5.3.1.6 Ovaries

As previously discussed, the volume of read counts produced from the samples decreased as the genital tract was ascended, with bacterial signal levels maxing out at 100 read counts in the ovarian samples. As a result, no firm conclusions could be drawn from the ovarian samples. Taking this into account,

reviewing the data that were produced at genera level showed that two thirds of both the benign (n/N 10/15, 66.6%) and malignant (n/N 10/16, 62.5%) samples were Lactobacillus-deplete (Fishers exact P=1) (Figure 5.6, Table 5.6).

With the same proviso regarding the volume of data, at species level the samples reflected the same findings as at genera level (Chi-squared p=0.699).

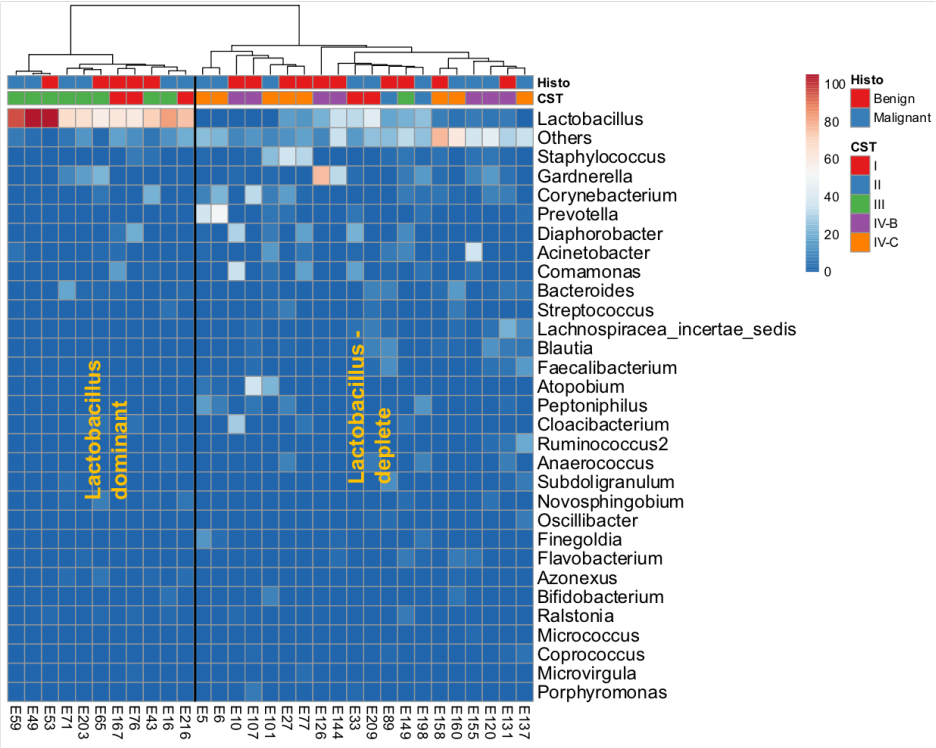


Figure 5.6 Hierarchal clustering of ovarian samples using Ward linkage and Euclidean distance

Table 5.6 Results of hierarchal clustering of ovarian samples using Euclidean distance and Ward linkage

Ovary	Lactobacillus- dominant n/N (%)	Lactobacillus deplete n/N (%)	Total n/N (%)
Benign	5/15 (33.3)	10/15 (66.6)	15/15 (100)
Cancer	6/16 (37.5)	10/16 (62.5)	15/15 (100)
Total	11/31 (35.5)	20/31 (64.5)	22/22 (100)
*Fisher' exact			P=1

5.3.2 Each anatomical site was noted to display distinguishing species between the benign and malignant samples

5.3.2.1 Vagina

Further study of the samples to determine if there was a compositional pattern was performed using univariate analysis. At species level, both *Blautia wexlerae* (P=0.0065214) and *Blautia faecis* (P=0.026435), were identified in higher relative abundance in the malignant samples (Figure 5.7). They did not withstand correction for false discovery rate (*B. wexlerae* P=0.5, *B. faecis* P= 0.5).

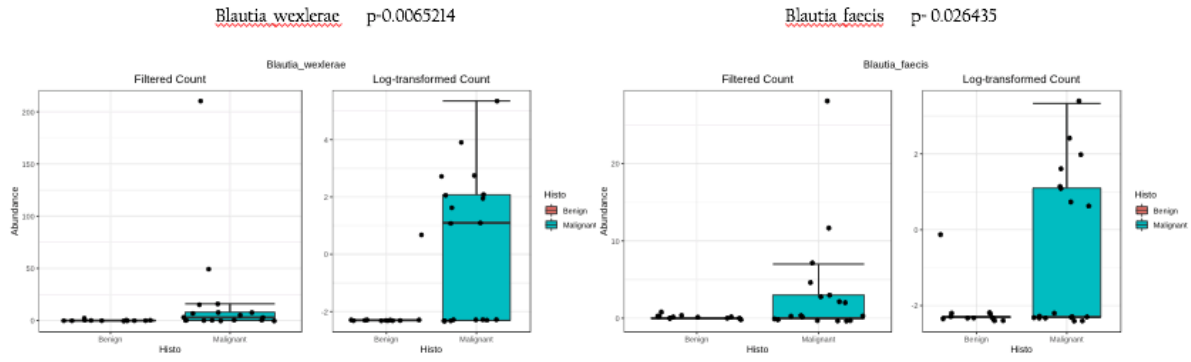


Figure 5.7 Univariate analysis of benign and malignant vaginal samples. The relative abundance of two bacterial species were found to be significantly higher in the malignant samples.

5.3.2.2 Cervix

The cervix yielded two further species that were found in differing abundance between the benign and malignant samples (Figure 5.8). The benign group had a higher relative abundance of *Dialistir micrarophilus* (P=0.019), whilst the malignant group had a higher relative abundance of *Lactobacillus fornicalis* (P=0.013). Similar to the vaginal samples, neither of these withstood correction for false discovery rate (*D. micrarophilus* P=0.7, *L.fornicalis* P=0.7) (Figure 5.8).

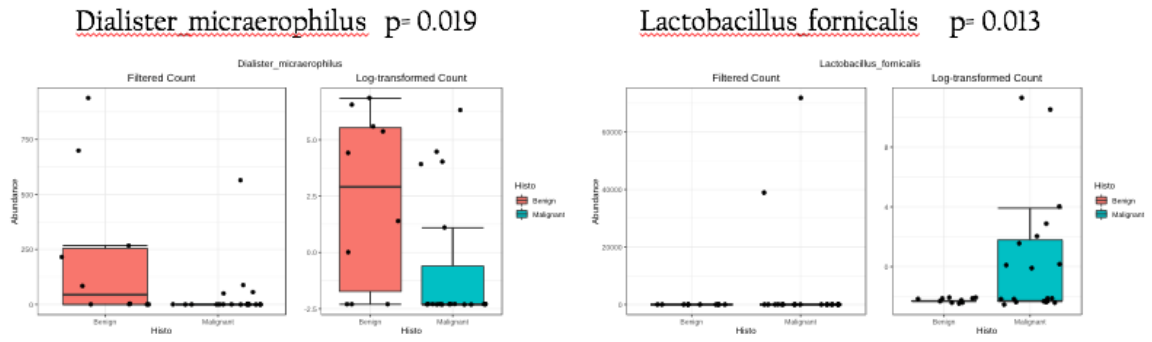


Figure 5.8 Univariate analysis of benign and malignant cervical samples. Dialister_micraerophilus was found in higher abundance in benign samples, whilst the malignant samples were enriched with Lactobacillus_fornicalis.

5.3.2.3 Endometrium

Both *Diaphorobacter aerolatus* (P=0.007) and *Dialister micraerophils* (P=0.007) were found in higher relative abundance in the benign samples without withstanding correction for false discovery rate (*D. aerolatus* P=0.42, *D. micraerophils* P= 0.42) (Figure 5.9).

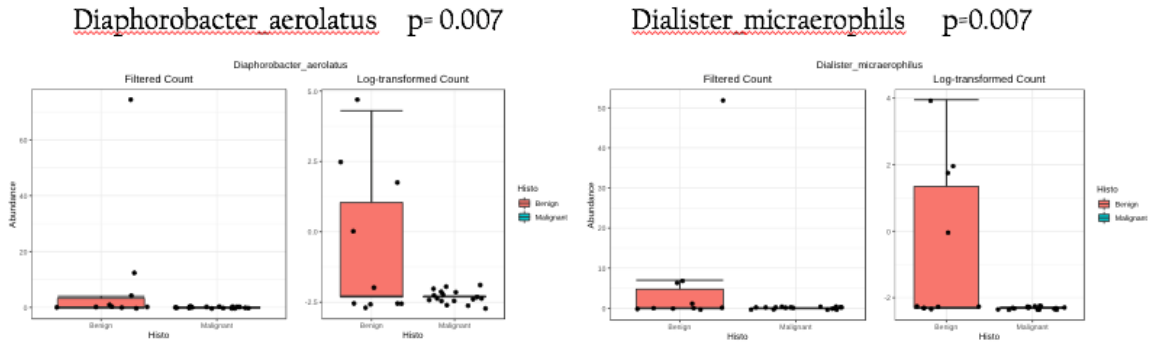


Figure 5.9 Univariate analysis of benign and malignant endometrial samples. Dialister aerolatus and Dialister micraerophilus were found in higher amounts in the benign samples.

5.3.2.4 Fallopian tubes

Facklamia hominis (P=0.039; FDR P=0.91) was present in higher amounts in the benign samples of the fallopian tubes, whilst *Lactobacillus fornicalis* (P=0.008; FDR P=0.91) was present in higher amounts in the malignant samples, a similar trend to the cervical malignant samples (Figure 5.10).

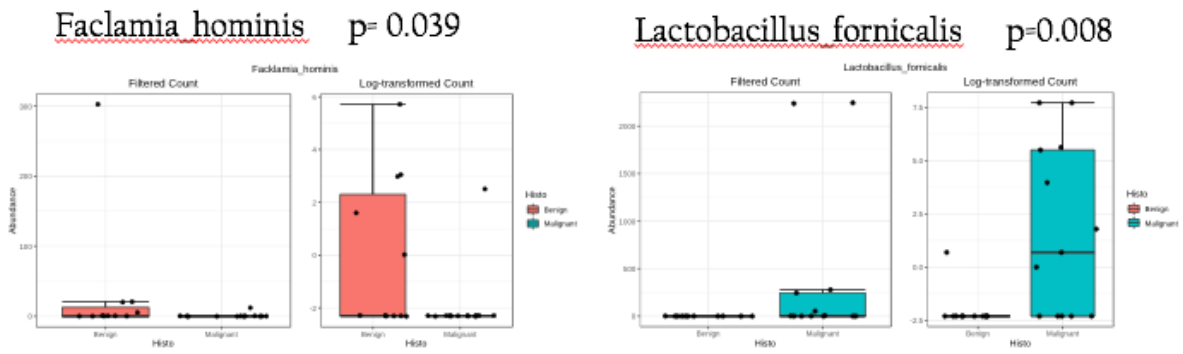


Figure 5.10 Univariate analysis of benign and malignant fallopian tube samples. Facklamia hominis was found in higher amounts in the benign samples whilst Lactobacillus fornicalis was found in higher amounts in the malignant samples.

5.3.2.5 Ovaries

At phylum level, there was an enrichment of Firmicutes (54%), followed by Actinobacteria (37%) with Proteobacteria at 5% in the benign cohort. The malignant samples had an even higher abundance of Firmicutes (71%), with Bacteroidetes the next prevalent (14%) followed by Proteobacteria (8%) and then Actinobacteria (6%) (Figure 5.11).

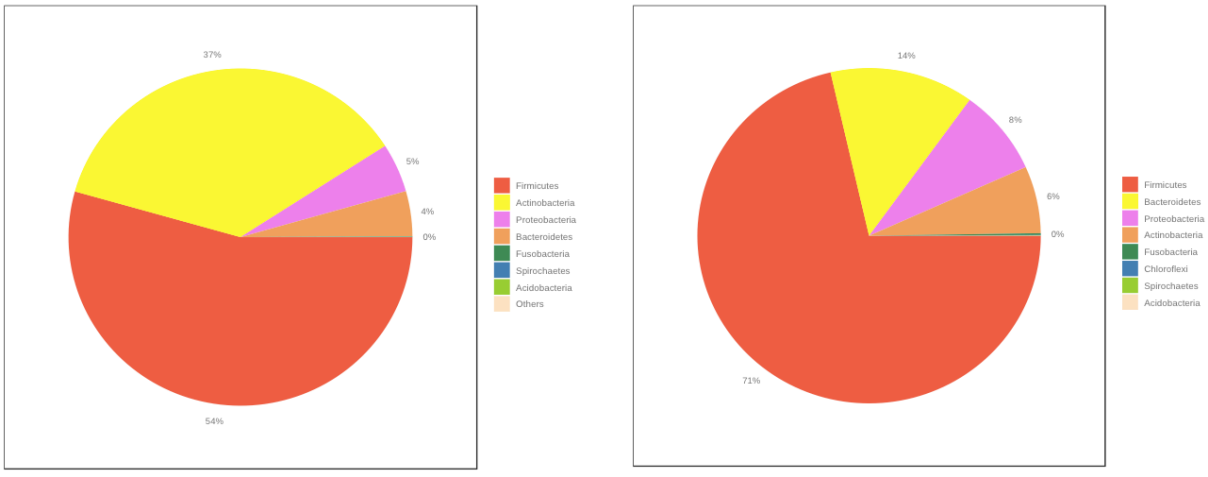


Figure 5.11 Proportion of bacteria in each cohort at phylum level.

Papillibacter unclassified (P=0.026) was seen to be present in higher relative abundance in the benign ovarian samples versus the malignant samples, which had higher relative abundance of *Dialister proprionicifaciens* (P=0.023). Neither species withstood correction for false discovery rate (*P. unclassified* P= 1, *D. proprionicifaciens* P=1) (Figure 5.12).

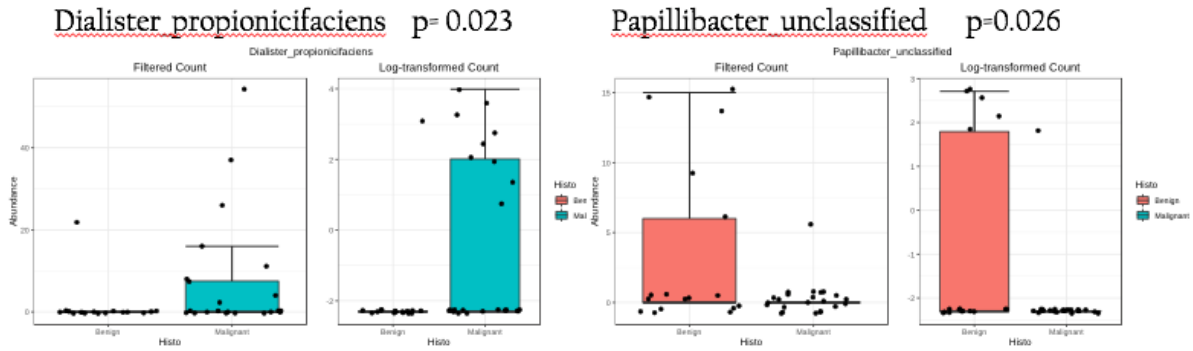


Figure 5.12 Univariate analysis of benign and malignant ovarian samples. Papillibacter unclassified was found in higher amounts in the benign samples whilst Dialister proprionicifaciens was found in higher amounts in the malignant samples.

5.3.3 Ovarian cancer samples displayed higher α-diversity than the benign controls at both species and genus level in the lower genital tract.

The α- and β-diversity of the community bacterial composition of both the ovarian cancer and benign controls samples were examined (Figure 5.13). Using Shannon a-diversity index and applying a Mann-Whitney/Kruskal-Wallis statistical test, both the vaginal (genus P=0.026, species P=0.018) and cervical (genus P=0.024, species P=0.008) ovarian cancer samples displayed a higher α-diversity than the benign controls at species level. β-diversity, or between sample diversity, when examined using Bray-Curtis index and a PERMANOVA statistical test, was not significant between the malignant and benign samples.

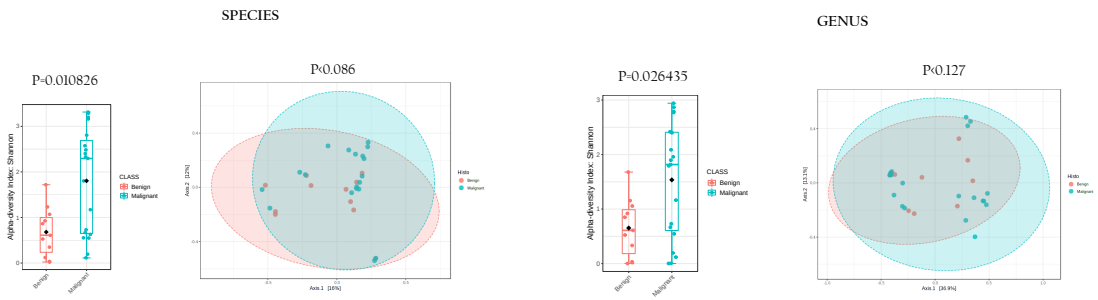


Figure 5.13 α and β diversity of vaginal samples at species and genus level. The malignant samples were statistically more diverse than the benign samples.

Ascending the genital tract, there was no difference in either diversity index between the ovarian cancer or benign controls groups in the endometrium, fallopian tubes nor the ovaries (see other results in Figure S3 in appendices).

5.3 DISCUSSION

In this chapter, I investigated whether different bacterial signatures could be identified between the benign and malignant cohorts and if specific signatures in the cervicovaginal niche were correlated with ovarian cancer. Whilst more evidence is emerging investigating the presence and composition of microbes in the female genital tract in relation to ovarian cancer, there is sparse literature examining the genital tract in its entirety, with most studies focusing on the upper genital tract. Whilst Qin et al. has presented the most complete study, it does not assess the vaginal microniche and overall is a broader study including a comparison of ovarian cancer tissue bacterial composition between those who had and had not undergone tubal ligation and those who were and were not resistant to platinum-based chemotherapy, with less focus on the specific composition of each anatomical site in each histological cohort.
As such, to my knowledge, this is the first study to assess the female genital tract in its entirety in relation to ovarian cancer with a particular focus on describing the composition of the microbial community structure in each anatomical area of the female genital tract.

The initial analysis approach used in this study implemented hierarchal clustering by Euclidean distance and Ward linkage, which enabled samples to be clustered into groups based on similarity of relative abundance of taxa. A common problem with this approach is that resulting clusters are determined by variance of profiles within this confined dataset and are therefore not broadly applicable to other datasets. This is a widely recognized issue in microbiome studies and to address this, the Ravel group recently developed the *VA*gina*L* community state typ*E* *N*earest Centro*I*d cl*A*ssifier or VALENCIA during data analysis[147]. Previous work by Ravel et al introduced the concept of CSTs to aid classification of the vaginal microbial structure[100]. The algorithm is based upon sequencing profiles of 13,160 vaginal samples thus allowing for improved capture of variance across diverse population thereby creating a standardized, reproducible clustering method. Whilst the VALENCIA method is based upon vaginal samples from reproductive age (mean age 30.6) women, the group also validated their results against several other cohorts including a postmenopausal group. As the VALENCIA method is optimized and developed for vaginal samples only, both clustering methods were employed in this study.

Interestingly, as is consistent with other published datasets, the cervicovaginal samples in the benign cohort were predominantly enriched with Lactobacillus and were low diversity. In contrast, high prevalence of Lactobacillus- deplete communities were observed in the malignant samples. This was demonstrated regardless of the clustering algorithm used. Similar results were seen for cervical samples. These findings are in line with results reported by Nené et al who described Lactobacillus-deplete cervicovaginal niche in both ovarian cancer and *BRCA* positive status[35]. Our group have demonstrated

a similar trend in a cervical cancer and CIN cohort, with increasing disease severity associated with decreasing relative abundance of *Lactobacillus spp*[104]. This adds to the hypothesis that vaginal dysbiosis is a potential hallmark for gynaecological malignancies. It is hypothesised that the loss of protective Lactobacillus, which maintains an inhospitable pH, produces species-specific metabolites, bacteriocins and competitively adheres to mucous, encourages a proinflammatory state that contributes to disease pathology.

A stark contrast is noted in the read counts as the UGT was ascended. The read counts dropped sharply from the cervix (1200 counts), to the endocervical junction (250 counts), with only the fallopian tubes displaying slightly higher counts (300 counts). Consequently, it was difficult to gain any firm conclusions from these results and they should be regarded with caution. Reviewing the numerical data however, the lower endometrial samples from the endocervical junction continued with a Lactobacillus-deplete signature in both the benign (n/N 9/14, 64.3%) and malignant samples (n/N 8/11, 72.7%). The fundus of the endometrium reflected the reverse however, with an enrichment of Lactobacillus or Gardnerella in 63% (n/N 7/11) of the benign samples and Lactobacillus- dominance noted in 73% (n/N 8/11) of the malignant samples. This contrasts with previous studies which have reported that the endometrial microniche is Lactobacillus- deplete in both benign conditions and other gynaecological cancers[38,107-110]. Whether this is a true finding in this cohort or are biased results due to the small sample size is difficult to assess. Another theory to consider as to why the fundus has a differing bacterial signature, is that the fundus is disturbed frequently secondary to hormonal changes and physical shedding during the menstrual cycle. The endocervical niche may be easier to re-colonise from the cervicovaginal site due to its proximity, whilst the fallopian tubes and ovaries do not shed and are left relatively undisturbed.

The fallopian tubes from patients with ovarian cancer returned to the Lactobacillus-deplete trend (73% deplete), and although the benign samples were Lactobacillus-dominant, a much smaller margin was noted than the LGT (55% Lactobacillus-dominant).

My next goal was to analyse the results for evidence of a distinctive bacterial signature between the benign and malignant samples. In the vagina, both *Blautia wexlarae* and *Blautia faecis*, of the phylum Firmicutes, were enriched in the malignant samples whilst *Lactobacillus fornicalis* was noted in the malignant cervical samples. Whilst both Nené et al. and Qin et al examined the cervicovaginal niche in women with ovarian cancer, neither identified a dominant bacterial species in their ovarian cancer subjects[35,38]. Nené commented on a community that was either more or less 50% Lactobacillus dominant whilst Qin stated that the proportion of Firmicutes was highest in the cervical samples relative to the other sampled sites. Blautia is still in its infancy, with *B. wexlarae* isolated in 2008 and *B. faecis* in 2013[285]. *B. wexlarae* is a dominant species in the intestine, and Blautia functional genes have been

shown to be enriched with translational, ribosomal structure and biogenesis which are frequently associated with environmental stresses e.g., oxidative stresses. Other strains have shown to exhibit oestrogenic effects through activation of oestrogen receptors[285]. The main physiological function of Blautia has been reported to be anti-inflammatory suggesting that their presence in the malignant vaginal samples could reflect an effort to minimise the inflammatory processes. Notably no *Blautia spp.* has been reported in the literature to be present in either benign or malignant ovarian samples, undermining the validity of our findings.

The cervical samples were noted to be enriched with *L. fornicalis*, though it is evident that this is largely driven by two outlier samples. As a result, whether it is a biologically relevant result is questionable, and further underlines the need to repeat similar experiments with larger cohorts to minimise the effect of individual results.

The endometrial cavity was noted to be enriched with *Diaphorobacter aerolatus* in the malignant cohort, with *Dialister micraeophilus* found in the benign cohort. *D. micraeophilus* has been isolated from amniotic fluid, pilonidal abscesses and blood cultures[286]. There is minimal literature on *D. aerolatus* and as it has only been isolated from the air, it is more than likely a contaminant in our study. Qin noted that *Salmonella spp* was enriched in the endometrial cavity of both the ovarian cancer and control subjects.

No consensus has been reached on the bacterial community of the fallopian tubes in benign conditions or ovarian cancer with several suggestions being quoted as the dominant species including *Ralstonia*, *Acidovorax*, *Lactococcus piscium*, *Shigella* and *Bacteroides*[34,37,105]. In this study, *Facklamia hominis* was notably enriched in the benign samples which was also noted by Chen et al[106]. The malignant samples were found to be enriched with *Lactobacillus fornicalis*. Despite analysing the fallopian tube of patients with ovarian cancer, Qin et al did not comment specifically on the microbial structure of the fallopian tubes, nor did Brewster et al[37,38].

As the few previous studies on assessing the microbes in ovarian cancer focused on the ovaries themselves, the data was also analysed at phylum level. Both cohorts were enriched by Firmicutes, with a higher abundance in the malignant group. In the malignant group, the next most abundant phyla was Bacteroidetes, followed by Proteobacteria and finally Actinobacteria. The benign group was secondarily enriched with Actinobacteria and lastly Proteobacteria. These results are in disagreement with both Zhou et al and Bannerjee et al who reported an increased Proteobacteria to Firmicutes ratio in their ovarian cancer samples[33,34]. At species level, the ovarian samples differed between the benign and cancer cohort with *Dialister proprionfaciens* and *Atopobium vaginae* enriched in the cancer group and *Papillibacter* enriched in the benign group. There is scant mention made in the literature relating to *D. proprionfaciens* but it has been isolated from wounds and abscesses[286].

Interestingly both Zhou et al. and Brewster et al. demonstrated a prevalence of *Actinebacter lwoffi* in ovarian samples from their cancer cohorts and *Ralstonia* in their benign cohort[34,287].

Lastly, I examined the richness and evenness of each anatomical site in patients with and without ovarian cancer. The cervicovaginal niche from the cancer cohort displayed a higher α diversity than their benign counterparts indicating that there was increased intra-sample bacterial diversity as is common is oncobiosis. No difference was noted in the β diversity, or inter-sample. No other statistically significant differences were noted for any diversity indices for any other sampled sites. These results are in contrast with Qin who commented that the β diversity was higher in their malignant samples but analysed their data enmasse without differentiating between body sites. Brewster demonstrated a significant difference in richness, evenness and diversity between the fallopian tube and fimbria, and the fimbria and ovary whilst Zhou reported that the cancer cohort had slightly decreased diversity indices relative to the control group. Whether detection of any notable differences is this study is because it was underpowered or is an accurate representation should be established in future work with a bigger cohort.

5.4 STRENGTHS AND LIMITATIONS

To date no other study has assessed the female genital tract in its entirety in relation to ovarian cancer. I investigated each anatomical niche in both the cancerous and benign cohorts specifically identifying a bacterial community structure, differing bacterial signatures between the cancer and benign groups and both the α and β diversity of each site in both groups. A limitation of this study is its cohort size which was underpowered and inhibited meaningful interpretation of its results. Longitudinal samples from the same cohort of women could help unravel some of the mechanisms underpinning carcinogenic transformation of precursor lesions including endometriosis and STIC lesions, and from high-risk conditions e.g., PID to ovarian cancer.

Current oncobiome and microbiome studies in other conditions have already advanced to investigating the mechanisms underpinning development of the disease and this is the next step for ovarian cancer. Further longitudinal work is required on larger cohorts with subclassification by histological subtype including taxonomic, metagenomic and inflammatory processes.

5.5 CONCLUSIONS

In conclusion, numerically the cervicovaginal niche is Lactobacillus-deplete in women with ovarian cancer, although not statistically significant. This trend continues throughout the genital tract, with an

exception at the fundus of the endometrium. Members of Blautia phylum were noted in the malignant vaginal samples however their relevance is yet to be determined as whilst they have been associated with infective conditions, they are also noted to be intestinal commensals. *Lactobacillus fornicalis* was identified in the malignant cervical samples and fallopian tubes samples which requires further examination as it has limited description in the literature and its role is yet to be elucidated. The cervicovaginal niche of the ovarian cancer patients was more diverse intra-individually which contrasts with the usual healthy cervicovaginal habitat which is commonly dominated by a handful of microbes. This phenomenon has been seen regularly in oncobiomes in other gynaecological cancers.

5.6 STATEMENT OF CONTRIBUTION

Dr Ann Smith performed the processing of 16S rRNA bacterial gene sequencing data in Mothur using the MiSeq SOP pipeline and assignment of OTU taxonomies in Chapters 4 and 5.

CHAPTER SIX: CONCLUSIONS

6.1 SUMMARY OF HYPOTHESIS

Ovarian cancer is the deadliest of all gynaecological cancers in the developed world. Whilst a significant genetic component accounts for up to 15% of cases, including *BRCA1/2* and Lynch syndrome, the process of carcinogenesis for the remainder is still unclear. The most common variant, HGSC, which accounts for up to 70% of all cases is now considered to originate from precursor lesions, STIC, in the fallopian tubes, whilst clear cell and endometrioid carcinomas are associated with the precursor condition of endometriosis. The process of transformation from the development of STIC and endometriotic deposits and cysts to ovarian cancer is yet to be uncovered. Epidemiological evidence has linked the pro-inflammatory conditions of PID, obesity and diabetes with ovarian cancer, whilst use of metformin and NSAIDS is considered to reduce risk. As a result of extensive investigation using NGS into the role of the human microbiome in other cancers including colorectal, pancreatic and lung cancer, the female genital tract microbiome and its part in gynaecological oncogenesis, particularly cervical and endometrial, has begun to be explored. Nevertheless, its role in the carcinogenesis of ovarian cancer is a novel concept still in its infancy.

The key hypothesis of this thesis was that not only is the upper genital tract not a sterile environment, but that there is a distinct bacterial signature specific to ovarian cancer and a vaginal oncobiome which is also indicative of ovarian cancer. The first aim of this thesis was therefore to evaluate the strength and validity of existing epidemiological evidence for associations between non-genetic risk factors, specifically proinflammatory conditions including PID and other related conditions, and the risk of developing ovarian cancer, which may have been prone to bias and required evaluation by robust statistical methods. The second aim was to establish the presence of a genuine bacterial signature above that of background contamination in the female upper genital tract. This is particularly important in the context of sites of low biomass. The bacterial presence in the endometrium, fallopian tubes and ovaries of women with benign or extra-uterine conditions was assessed alongside the same anatomical sites in women with ovarian cancer. The next step was to investigate the theory of how microbes may have gained access to the UGT and whether a continuum or 'migration' from the LGT is a feasible consideration. Finally, this work aimed to characterise and compare the bacterial signature of the entire female genital tract in a benign cohort and cancer cohort and identify if a specific signature was present in the cervicovaginal niche.

6.2 SUMMARY OF FINDINGS

6.2.1 Umbrella review

Key findings

- ◊ Raised BMI, greater height and use of HRT increased the risk of ovarian cancer; use of OCP confers a reduction in the risk (strongly suggestive)
- ◊ Endometriosis increased the risk whilst use of metformin decreased the risk (highly suggestive)
- ◊ History of diabetes increased the risk of ovarian cancer (suggestive)
- ◊ PID and IVF increased the risk, and use of NSAIDS reduced the risk (weak)

An umbrella review of the available evidence for risk factors associated with ovarian cancer revealed that raised BMI of several different measurements was associated with an increase in ovarian cancer across all grades of evidence. Several other pro-inflammatory conditions were also associated with an increase in risk including endometriosis, diabetes and PID. The mechanisms that underlie the association of these risk factors with ovarian cancer are not fully characterised but feasibly include higher oestrogen levels in postmenopausal women, both systemic and localised dysbiosis and a chronic low-grade inflammatory state.

Conversely, use of the OCP, metformin and NSAIDS all conferred a reduction in risk. The OCP is thought to reduce the risk by minimising the number of ovulatory cycles whilst metformin and NSAIDS both have anti-inflammatory actions.

6.2.2 Female genital tract presence

Key findings

- ◊ A considerable proportion of sequencing data are contaminants, which is particularly influential in low biomass sites
- ◊ A biological signature is present in the female upper genital tract which is compositionally distinct from background contaminants
- ◊ All patients in the benign cohort were noted to harbour an uninterrupted continuum of vaginal bacterial species throughout the entire genital tract, with less robust findings in the cancer cohort (80% of patients demonstrate a continuum)
- ◊ The endometrium and UGT shared more bacterial species than the LGT and the endometrium; the fallopian tubes and ovaries had a more similar bacterial signature than the vagina and cervix.

Systematic and rigorous implementation of technical controls including environmental, equipment and kit controls in conjunction with stringent decontamination analyses revealed that 20.6% of sequencing was background DNA contaminants and subsequently removed. The presence of a bacterial signature in the endometrium, fallopian tubes and ovaries above background was noted in both cohorts, with fewer sequencing reads noted in the endometrium. All benign patient in the cohort demonstrated a continuum of bacterial signature at all anatomical sites, which were noted to be the most abundant vaginal species, lending weight to the concept of migration of bacterial from the LGT to the UGT. 80% of the cancer cohort showed the same phenomena, with 40% of the cohort indicating a vaginal species.

The LGT and UGT had a shared commonality of 45.45% of bacterial species. The endometrium and UGT had a higher degree of commonality than the LGT and the endometrium with consideration being given to the role of the cervical barrier and the repeated shedding of the endometrium throughout a female reproductive lifespan facilitating a less disturbed bacterial niche.

6.2.3 Female genital tract composition

Key findings

- ◊ Lactobacillus-depletion in the LGT and UGT, but not the endometrial cavity, is noted in women with ovarian cancer
- ◊ Distinguishing bacterial species were identified at each anatomical site (*Blautia wexlerae* and *Blautia faecis* in the vagina, *Papillibacter unclassified* and *Dialister proprionicifaciens* in the ovaries); their function is yet to be established
- ◊ Higher intra-individual diversity is noted at species and genus level in the LGT of the ovarian cancer cohort

Higher prevalence of high-diversity, Lactobacillus depleted communities were consistently observed in the ovarian cancer cohort, regardless of the clustering method used. Furthermore, Lactobacillus-depletion and high diversity was noted in the fallopian tubes and ovaries, albeit not significantly and to a lesser extent than the LGT. These findings are consistent both with the few studies to date on ovarian cancer, and with the current literature on both cervical and endometrial cancer. Whether Lactobacillus- depletion is an oncobiome and promotes a proinflammatory environment and subsequently carcinogenesis or is a consequence of carcinogenesis requires further investigation. These findings build upon the theory that the bacterial landscape and its pro or anti-inflammatory effects in the female genital tract requires stringent investigation and the underlying mechanisms could provide valuable cues for identification of early diagnostic biomarker, preventive and therapeutic interventions against ovarian cancer.

6.3 PROPOSED MECHANISM

Ovarian cancer is associated with several proinflammatory precursor states including PID, endometriosis, diabetes and obesity. The most common, and fatal, of ovarian cancer subtypes, HGSC, in those without genetic predisposition, is considered to arise primarily from STIC lesions within the fallopian tubes. The most common genetic mutation is *TP53* which encodes for the tumour suppressor protein, p53, and is found in 92% of STIC lesions[288,289] . Epidemiological evidence prior to this work reports that tubal ligation diminishes the risk of developing ovarian cancer.

One of the core processes in the development of malignancies is inflammation. It has been shown to be present both in the tissue surrounding neoplastic tumours and also within the tumours themselves[83]. Hallmarks of chronic or 'smouldering' inflammation in the tumour microenvironment which promote tumour growth include the presence of chemokines, cytokines and prostaglandins, tissue remodelling, neoangiogenesis and tissue repair[84]. These inflammatory-like processes aid in the proliferation of malignant cells, promotes metastasis, subverts adaptive immune responses and alters responses to hormones and chemotherapeutic agents[84].

Toll like receptors (TLRs) are pattern recognition receptors (PRRs) that sense danger by detecting non-self or altered molecular patterns such as pathogen-associated molecular patterns (PAMPs)[290]. PAMP-induced TLR activation triggers a profound activation of multiple proinflammatory and anti-viral signaling cascades to achieve pathogen elimination. If pathogens cannot be eliminated, however, they may elicit chronic inflammation, as seen in cases of PID. Indeed, chronic salpingitis has been found in 53% of ovarian carcinoma cases, with 80% of another cohort of ovarian cancer cases positive for *C. trachomatis*[32,291]. TLRs induce the important tumour-promoting signaling pathway for the transcription factor NF-kB. NF-kB is a master switch of inflammation regulating the transcription of >100 proinflammatory genes. TLR signaling upregulates well-known tumour promoting inflammatory cytokines through NF-kB-dependent pathways, including interleukin (IL)-1β, tumour necrosis factor a (TNFα) and IL-6. TLR2-mediated inflammatory signals in macrophages, triggered by tumour-derived, TLR2- agonistic extracellular matrix (ECM) protein versican, promote the secretion of TNFα and metastasis. NF-kB activation also results in a wide range of other cellular responses such as prevention of apoptosis, proliferation and antioxidant defenses, and importantly, has been reported to suppress the function of p53, the 'guardian of the genome' tumour suppressor protein which is most commonly mutated in ovarian cancer. Transcriptome sequencing (RNA-seq) analyses by Zhou et al. demonstrated upregulation of inflammation- associated signalling pathways including NK-κB, cytokine-cytokine receptor interaction and chemokine signaling pathways in ovarian cancer tissue samples[34]. They also reported that genes for TLR6, MyD88, NLRP3, inflammasomes which are involved

in activating the immune response, were increased in the ovarian cancer cohort[34]. *Acinetobacter baumanni*, of the same genus as *Acinetobacter guangdongensis* found in the both the benign and malignant ovarian samples, has been found to secrete lipopolysaccharides (LPS) and extracellular vesicles that activated TLR4 signaling cascade. Whole UV-killed cells of *A. baumannii* and *Acinetobacter* 'genomospecies 9' stimulated both TLR2- and TLR4-dependent signalling[292].

In the FGT, bacterial vaginosis (BV), which is known to create biofilms, is characterised by a dysbiotic microbiome with a lower prevalence of lactobacilli and a higher prevalence of anaerobic bacteria. It has been shown to be associated with several adverse health outcomes including sexually transmitted diseases including HPV and HIV, PID and endometritis[293-296]. BV biofilms have been noted to translocate to the UGT, feasibly as a result of the peristaltic activity of the uterus and fallopian tubes[284]. They have also been found in areas of low biomass without endogenous microbes[297]. Furthermore, it has also been demonstrated that even if the microbial inhabitants within the biofilm are dead, they can still elicit an inflammatory response, as reported by Mandal et al.[292]. Consequently, this suggests that it is feasible that an acute bacterial insult may occur in younger years e.g., PID, with carcinogenic sequalae not occurring until postmenopausally, potentially because of the hypo-oestrogenic state leading to insufficient glycogen levels and subsequent low levels of health- promoting lactobacilli. These proinflammatory pathways described above may intersect with dysbiosis of the female genital tract i.e., BV or PID, inducing carcinogenesis. A eubiotic state may promote immune activation, prevent pathogen proliferation by competitive presence, produce bactericidal molecules and in the vagina, maintain a hostile acidic environment as orchestrated by elevated levels of lactobacilli, thus hampering the ascension of pathogenic bacteria. A dysbiotic state may trigger an acute inflammatory reaction as described above, which leads to a chronic inflammatory reaction, tissue remodelling, and the subsequent consequences of uninhibited inflammatory pathways and tumorigenesis as seen in conditions such as gastric cancer (*H. pylori*), hepatic carcinoma (hepatitis), and cholangiocarcinoma (*S. enterica Typhi*).

6.4 TRANSLATIONAL VALUE

Despite increased public awareness of ovarian cancer and significant advances in the treatment of ovarian cancer including PARP inhibitors and HIPEC, it still remains the deadliest gynaecological malignancy in the developed world. The multi-factorial aetiology of ovarian cancer is complex and the signalling pathways and genetic mutations that are activated leading to ovarian cancer development need further investigation. The ability to identify women who are at high-risk of developing ovarian cancer, particularly those that have completed their families or are postmenopausal, may change clinical practice in the future through recommending early bilateral salpingectomy to eliminate the risk of HGSC developing in the future. The microbiome has been demonstrated to affect the effectiveness of

chemotherapy and immunotherapy for several cancers e.g., melanoma, where patients who received antibiotics prior to administration of immunotherapy experienced decreased effectiveness and survival time. Targeting the microbiome for manipulation has already seen success in irritable bowel syndrome patients with recurrent *Clostridium difficile* infection, where a faecal microbiota transplantation returned gut eubiosis. An exploratory trial investigating topical vaginal application of a probiotic *Lactobacillus spp.* in patients with recurrent BV demonstrated that colonization was achieved in 44% of participants and inhibited the growth of BV- related species[272].

6.5 FUTURE DIRECTION

Further studies to thoroughly characterise the female genital tract in health and cancerous states are warranted to refine which microbes or microbial signatures have the potential for manipulation.

Epidemiological studies could encompass:

- Case- control studies of the female genital tract of cohorts with varied racial and ethnic backgrounds, reproductive factors e.g., tubal ligation, menopausal status, diet and BMI, and socio-economic factors.

Longitudinal studies on the microbiome should have larger, multicentre cohorts to achieve more reliable, well powered results. They should investigate:

- The effect of chemotherapy and immunotherapy on the FGT and the gut microbiome including use of pre/pro and antibiotics on the effectiveness of therapies and management of immune-related adverse events (IrAEs).
- The LGT of women at risk of ovarian cancer whether due to genetic factors e.g., BRCA mutations, or history of endometriosis or PID to monitor any precancerous changes which may suggest precursor biomarkers of disease
- Any differences depending on histological subtype or grade of disease
- Integrated metagenomic, metabolomic, metatranscriptonomic, metaproteomic and microbiome studies of these women with ovarian cancer would provide invaluable information on the host-microbiome interactions which would further help to personalize therapeutics.

Laboratory-based organoid models of ovarian cancer to test the effects of specific microbes e.g., *C trachomatis* and *Mycoplasma spp.*, both of which have been found in several studies in ovarian cancer tissue, and 'dysbiotic' communities should be undertaken, with systematic and thorough evaluation of inflammatory signalling pathways.

As has been discussed in this work, contaminant concerns are at the forefront of low biomass studies and future microbiome studies should be validated by undergoing repeated and different extraction and detection methods e.g., 16S rRNA, amplicon sequencing variants (ASVs) along with OTUs etc., to optimise the likeliness of a 'true biological signature'. Propidium monoazide (PMA), a DNA-binding dye, should also be used to discriminate between living and dead cells in these low biomass sites, and all eukaryotes should be assessed.

REFERENCES

1. Kurman RJ, Shih Ie M. The origin and pathogenesis of epithelial ovarian cancer: a proposed unifying theory. *Am J Surg Pathol* 2010; **34**(3): 433-43.
2. UK CR. Ovarian cancer statistics. 2023. https://crukcancerintelligence.shinyapps.io/EarlyDiagnosis/, (accessed November 2023).
3. 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**(3): 209-49.
4. Cancer IAfRo. Cancer Today. 2020. https://gco.iarc.fr/today/online-analysis-table?v=2020&mode=cancer&mode_population=continents&population=900&populations=908_900_826&key=asr&sex=2&cancer=39&type=0&statistic=5&prevalence=0&population_group=0&ages_group%5B%5D=0&ages_group%5B%5D=17&group_cancer=1&include_nmsc=0&include_nmsc_other=1#collapse-group-0-4 (accessed November 2023).
5. Prat J, Oncology FCoG. Staging Classification for Cancer of the Ovary, Fallopian Tube, and Peritoneum: Abridged Republication of Guidelines From the International Federation of Gynecology and Obstetrics (FIGO). *Obstet Gynecol* 2015; **126**(1): 171-4.
6. Berek JS, Renz M, Kehoe S, Kumar L, Friedlander M. Cancer of the ovary, fallopian tube, and peritoneum: 2021 update. *Int J Gynaecol Obstet* 2021; **155 Suppl 1**: 61-85.
7. Vang R, Shih IM, Kurman RJ. Fallopian tube precursors of ovarian low- and high-grade serous neoplasms. *Histopathology* 2013; **62**(1): 44-58.
8. Kroeger PT, Jr., Drapkin R. Pathogenesis and heterogeneity of ovarian cancer. *Curr Opin Obstet Gynecol* 2017; **29**(1): 26-34.
9. Young RH, Carcangiu ML, Herrington CS, Kurman RJ. WHO classification of tumours of female reproductive organs. Fourth edition. ed. Lyon, [France: International Agency for Research on Cancer; 2014.
10. Kim J, Park EY, Kim O, et al. Cell Origins of High-Grade Serous Ovarian Cancer. *Cancers (Basel)* 2018; **10**(11).
11. Kuhn E, Meeker A, Wang TL, Sehdev AS, Kurman RJ, Shih Ie M. Shortened telomeres in serous tubal intraepithelial carcinoma: an early event in ovarian high-grade serous carcinogenesis. *Am J Surg Pathol* 2010; **34**(6): 829-36.
12. WCRF/AICR. The Associations between Food, Nutrition, and Physical Activity and the Risk of Ovarian Cancer Continuous Update Project Report. *WCRF/ AICR* 2013.
13. Pietragalla A, Arcieri M, Marchetti C, Scambia G, Fagotti A. Ovarian cancer predisposition beyond BRCA1 and BRCA2 genes. *Int J Gynecol Cancer* 2020; **30**(11): 1803-10.
14. Neff RT, Senter L, Salani R. BRCA mutation in ovarian cancer: testing, implications and treatment considerations. *Ther Adv Med Oncol* 2017; **9**(8): 519-31.
15. Brown KF, Rumgay H, Dunlop C, et al. The fraction of cancer attributable to modifiable risk factors in England, Wales, Scotland, Northern Ireland, and the United Kingdom in 2015. *Br J Cancer* 2018; **118**(8): 1130-41.
16. Wild C, Weiderpass, E., Stewart, BW. World Cancer Report: Cancer Research for Cancer Prevention. Lyon, France: International Agency for Research on Cancer, 2020.
17. Reid A, de Klerk N, Musk AW. Does exposure to asbestos cause ovarian cancer? A systematic literature review and meta-analysis. *Cancer Epidemiol Biomarkers Prev* 2011; **20**(7): 1287-95.
18. Langseth H, Johansen BV, Nesland JM, Kjaerheim K. Asbestos fibers in ovarian tissue from Norwegian pulp and paper workers. *Int J Gynecol Cancer* 2007; **17**(1): 44-9.
19. Benyi E, Linder M, Adami J, Kieler H, Palme M, Savendahl L. Adult height is associated with risk of cancer and mortality in 5.5 million Swedish women and men. *J Epidemiol Community Health* 2019; **73**(8): 730-6.
20. Wild C, Weiderpass, E., Stewart, BW. World Cancer Report; Cancer Research for Cancer Prevention. In: Wild C, Weiderpass, E., Stewart, BW., editor. Lyon, France: International Agency for Research on Cancer; 2020.

21. Sainz de la Cuesta R, Eichhorn JH, Rice LW, Fuller AF, Jr., Nikrui N, Goff BA. Histologic transformation of benign endometriosis to early epithelial ovarian cancer. *Gynecol Oncol* 1996; **60**(2): 238-44.
22. Komiyama S, Aoki D, Tominaga E, Susumu N, Udagawa Y, Nozawa S. Prognosis of Japanese patients with ovarian clear cell carcinoma associated with pelvic endometriosis: clinicopathologic evaluation. *Gynecol Oncol* 1999; **72**(3): 342-6.
23. Shu XO, Brinton LA, Gao YT, Yuan JM. Population-based case-control study of ovarian cancer in Shanghai. *Cancer Res* 1989; **49**(13): 3670-4.
24. Zhou Z, Zeng F, Yuan J, et al. Pelvic inflammatory disease and the risk of ovarian cancer: a meta-analysis. *Cancer Causes Control* 2017; **28**(5): 415-28.
25. Piao J, Lee EJ, Lee M. Association between pelvic inflammatory disease and risk of ovarian cancer: An updated meta-analysis. *Gynecol Oncol* 2020; **157**(2): 542-8.
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**(1): 55-67.
27. Wang C, Liang Z, Liu X, Zhang Q, Li S. The Association between Endometriosis, Tubal Ligation, Hysterectomy and Epithelial Ovarian Cancer: Meta-Analyses. *Int J Environ Res Public Health* 2016; **13**(11).
28. Chan PJ, Seraj IM, Kalugdan TH, King A. Prevalence of mycoplasma conserved DNA in malignant ovarian cancer detected using sensitive PCR-ELISA. *Gynecol Oncol* 1996; **63**(2): 258-60.
29. Idahl A, Lundin E, Jurstrand M, et al. Chlamydia trachomatis and Mycoplasma genitalium plasma antibodies in relation to epithelial ovarian tumors. *Infect Dis Obstet Gynecol* 2011; **2011**: 824627.
30. Paradowska E, Jablonska A, Studzinska M, Wilczynski M, Wilczynski JR. Detection and genotyping of CMV and HPV in tumors and fallopian tubes from epithelial ovarian cancer patients. *Sci Rep* 2019; **9**(1): 19935.
31. Idahl A, Le Cornet C, Gonzalez Maldonado S, et al. Serologic markers of Chlamydia trachomatis and other sexually transmitted infections and subsequent ovarian cancer risk: Results from the EPIC cohort. *Int J Cancer* 2020; **147**(8): 2042-52.
32. Shanmughapriya S, Senthilkumar G, Vinodhini K, Das BC, Vasanthi N, Natarajaseenivasan K. Viral and bacterial aetiologies of epithelial ovarian cancer. *Eur J Clin Microbiol Infect Dis* 2012; **31**(9): 2311-7.
33. Banerjee S, Tian T, Wei Z, et al. The ovarian cancer oncobiome. *Oncotarget* 2017; **8**(22): 36225-45.
34. Zhou B, Sun C, Huang J, et al. The biodiversity Composition of Microbiome in Ovarian Carcinoma Patients. *Sci Rep* 2019; **9**(1): 1691.
35. Nené 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**(8): 1171-82.
36. Wang Q, Zhao L, Han L, et al. The differential distribution of bacteria between cancerous and noncancerous ovarian tissues in situ. *J Ovarian Res* 2020; **13**(1): 8.
37. Brewster WR, Burkett WC, Ko EM, Bae-Jump V, Nicole McCoy A, Keku TO. An evaluation of the microbiota of the upper reproductive tract of women with and without epithelial ovarian cancer. *Gynecol Oncol Rep* 2022; **42**: 101017.
38. Qin X, Zhou J, Wang Z, et al. Metagenomic analysis of the microbiome of the upper reproductive tract: combating ovarian cancer through predictive, preventive, and personalized medicine. *EPMA J* 2022; **13**(3): 487-98.
39. Risch HA, Howe GR. Pelvic inflammatory disease and the risk of epithelial ovarian cancer. *Cancer Epidemiol Biomarkers Prev* 1995; **4**(5): 447-51.
40. Parazzini F, La Vecchia C, Negri E, Moroni S, dal Pino D, Fedele L. Pelvic inflammatory disease and risk of ovarian cancer. *Cancer Epidemiol Biomarkers Prev* 1996; **5**(8): 667-9.
41. Ness RB, Grisso JA, Cottreau C, et al. Factors related to inflammation of the ovarian epithelium and risk of ovarian cancer. *Epidemiology* 2000; **11**(2): 111-7.

42. Merritt MA, Green AC, Nagle CM, Webb PM, Australian Cancer S, Australian Ovarian Cancer Study G. Talcum powder, chronic pelvic inflammation and NSAIDs in relation to risk of epithelial ovarian cancer. *Int J Cancer* 2008; **122**(1): 170-6.
43. Lin HW, Tu YY, Lin SY, et al. Risk of ovarian cancer in women with pelvic inflammatory disease: a population-based study. *Lancet Oncol* 2011; **12**(9): 900-4.
44. Christina BR, Mette TF, Allan J, et al. Pelvic inflammatory disease and risk of invasive ovarian cancer and ovarian borderline tumors. *Cancer Causes & Control* 2013; **24**(7): 1459-64.
45. Rasmussen CB, Kjaer SK, Albieri V, et al. Pelvic Inflammatory Disease and the Risk of Ovarian Cancer and Borderline Ovarian Tumors: A Pooled Analysis of 13 Case-Control Studies. *Am J Epidemiol* 2017; **185**(1): 8-20.
46. Falconer H, Yin L, Salehi S, Altman D. Association between pelvic inflammatory disease and subsequent salpingectomy on the risk for ovarian cancer. *Eur J Cancer* 2021; **145**: 38-43.
47. Perri T, Levin G, Naor-Revel S, et al. Risk-reducing salpingo-oophorectomy and breast cancer incidence among Jewish BRCA1/BRCA2-mutation carriers-an Israeli matched-pair study. *Int J Gynaecol Obstet* 2022; **157**(2): 431-6.
48. Choi YH, Terry MB, Daly MB, et al. Association of Risk-Reducing Salpingo-Oophorectomy With Breast Cancer Risk in Women With BRCA1 and BRCA2 Pathogenic Variants. *JAMA Oncol* 2021; **7**(4): 585-92.
49. Hanley GE, Kwon JS, Finlayson SJ, Huntsman DG, Miller D, McAlpine JN. Extending the safety evidence for opportunistic salpingectomy in prevention of ovarian cancer: a cohort study from British Columbia, Canada. *Am J Obstet Gynecol* 2018; **219**(2): 172 e1- e8.
50. Hanley GE, Kwon JS, McAlpine JN, Huntsman DG, Finlayson SJ, Miller D. Examining indicators of early menopause following opportunistic salpingectomy: a cohort study from British Columbia, Canada. *Am J Obstet Gynecol* 2020; **223**(2): 221 e1- e11.
51. Kwon JS, McAlpine JN, Hanley GE, et al. Costs and benefits of opportunistic salpingectomy as an ovarian cancer prevention strategy. *Obstet Gynecol* 2015; **125**(2): 338-45.
52. McAlpine JN, Hanley GE, Woo MM, et al. Opportunistic salpingectomy: uptake, risks, and complications of a regional initiative for ovarian cancer prevention. *Am J Obstet Gynecol* 2014; **210**(5): 471 e1-11.
53. Azam S, Lange T, Huynh S, et al. Hormone replacement therapy, mammographic density, and breast cancer risk: a cohort study. *Cancer Causes Control* 2018; **29**(6): 495-505.
54. Buys SS, Partridge E, Black A, et al. Effect of screening on ovarian cancer mortality: the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Randomized Controlled Trial. *JAMA* 2011; **305**(22): 2295-303.
55. Jacobs IJ, Menon U, Ryan A, et al. Ovarian cancer screening and mortality in the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS): a randomised controlled trial. *Lancet* 2016; **387**(10022): 945-56.
56. Pinsky PF, Yu K, Kramer BS, et al. Extended mortality results for ovarian cancer screening in the PLCO trial with median 15years follow-up. *Gynecol Oncol* 2016; **143**(2): 270-5.
57. 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**(10290): 2182-93.
58. Tamahsebi F. Raised CA125 - what we actually know.... 2020.
59. Silberstein LB, Rosenthal AN, Coppack SW, Noonan K, Jacobs IJ. Ascites and a raised serum Ca 125--confusing combination. *J R Soc Med* 2001; **94**(11): 581-2.
60. Zuckerman E, Lanir A, Sabo E, et al. Cancer antigen 125: a sensitive marker of ascites in patients with liver cirrhosis. *Am J Gastroenterol* 1999; **94**(6): 1613-8.
61. Moley KH, Massad LS, Mutch DG. Pelvic inflammatory disease. Correlation of severity with CA-125 levels. *J Reprod Med* 1996; **41**(5): 341-6.
62. Karimi-Zarchi M, Dehshiri-Zadeh N, Sekhavat L, Nosouhi F. Correlation of CA-125 serum level and clinico-pathological characteristic of patients with endometriosis. *Int J Reprod Biomed* 2016; **14**(11): 713-8.

63. Huang F, Chen J, Liu Y, Zhang K, Wang J, Huang H. New mechanism of elevated CA125 in heart failure: the mechanical stress and inflammatory stimuli initiate CA125 synthesis. *Med Hypotheses* 2012; **79**(3): 381-3.
64. Yanaranop M, Jantarateptewan N, Tiyayon J, Nakrangsee S. Significance of Serum Human Epididymis Protein 4 and Cancer Antigen 125 in Distinguishing Type I and Type II Epithelial Ovarian Cancers. *Int J Gynecol Cancer* 2018; **28**(6): 1058-65.
65. Topalak O, Saygili U, Soyturk M, et al. Serum, pleural effusion, and ascites CA-125 levels in ovarian cancer and nonovarian benign and malignant diseases: a comparative study. *Gynecol Oncol* 2002; **85**(1): 108-13.
66. Ovarian Cancer: Recognition and initial management (CG122). 2011.
67. Nunes N, Ambler G, Foo X, Naftalin J, Widschwendter M, Jurkovic D. Use of IOTA simple rules for diagnosis of ovarian cancer: meta-analysis. *Ultrasound Obstet Gynecol* 2014; **44**(5): 503-14.
68. Kaijser J, Bourne T, De Rijdt S, et al. Key findings from the International Ovarian Tumor Analysis (IOTA) study: an approach to the optimal ultrasound based characterisation of adnexal pathology. *Australas J Ultrasound Med* 2012; **15**(3): 82-6.
69. Kurman RJ, Shih Ie M. The Dualistic Model of Ovarian Carcinogenesis: Revisited, Revised, and Expanded. *Am J Pathol* 2016; **186**(4): 733-47.
70. Jayson GC, Kohn EC, Kitchener HC, Ledermann JA. Ovarian cancer. *The Lancet*; **384**(9951): 1376-88.
71. Harbin LM, Gallion HH, Allison DB, Kolesar JM. Next Generation Sequencing and Molecular Biomarkers in Ovarian Cancer-An Opportunity for Targeted Therapy. *Diagnostics (Basel)* 2022; **12**(4).
72. Prat J. Pathology of borderline and invasive cancers. *Best Pract Res Clin Obstet Gynaecol* 2017; **41**: 15-30.
73. Prat J, D'Angelo E, Espinosa I. Ovarian carcinomas: at least five different diseases with distinct histological features and molecular genetics. *Hum Pathol* 2018; **80**: 11-27.
74. Crum CP, Drapkin R, Kindelberger D, Medeiros F, Miron A, Lee Y. Lessons from BRCA: the tubal fimbria emerges as an origin for pelvic serous cancer. *Clin Med Res* 2007; **5**(1): 35-44.
75. Kindelberger DW, Lee Y, Miron A, et al. Intraepithelial carcinoma of the fimbria and pelvic serous carcinoma: Evidence for a causal relationship. *Am J Surg Pathol* 2007; **31**(2): 161-9.
76. Zafrakas M, Grimbizis G, Timologou A, Tarlatzis BC. Endometriosis and ovarian cancer risk: a systematic review of epidemiological studies. *Front Surg* 2014; **1**: 14.
77. Sayasneh A, Tsivos D, Crawford R. Endometriosis and ovarian cancer: a systematic review. *ISRN Obstet Gynecol* 2011; **2011**: 140310.
78. Brilhante AV, Augusto KL, Portela MC, et al. Endometriosis and Ovarian Cancer: an Integrative Review (Endometriosis and Ovarian Cancer). *Asian Pac J Cancer Prev* 2017; **18**(1): 11-6.
79. Kirstine KS, Michelle ZL, Kristina L, Dóra KF, Margaret RK, Henrik TS. Pelvic inflammatory disease and risk of cancer: A nationwide cohort study. *Int J Gynecol Obstet* 2019.
80. Tanaka Y, Mori T, Ito F, et al. Exacerbation of Endometriosis Due To Regulatory T-Cell Dysfunction. *J Clin Endocrinol Metab* 2017; **102**(9): 3206-17.
81. Liu F, He L, Liu Y, Shi Y, Du H. The expression and role of oxidative stress markers in the serum and follicular fluid of patients with endometriosis. *Clin Exp Obstet Gynecol* 2013; **40**(3): 372-6.
82. Dzutsev A, Goldszmid RS, Viaud S, Zitvogel L, Trinchieri G. The role of the microbiota in inflammation, carcinogenesis, and cancer therapy. *Eur J Immunol* 2015; **45**(1): 17-31.
83. Balkwill F, Mantovani A. Inflammation and cancer: back to Virchow? *Lancet* 2001; **357**.
84. Mantovani A, Allavena P, Sica A, Balkwill F. Cancer-related inflammation. *Nature* 2008; **454**(7203): 436-44.
85. Madeleine MM, Anttila T, Schwartz SM, et al. Risk of cervical cancer associated with Chlamydia trachomatis antibodies by histology, HPV type and HPV cofactors. *Int J Cancer* 2007; **120**(3): 650-5.
86. Vergote I, Trope CG, Amant F, et al. Neoadjuvant chemotherapy or primary surgery in stage IIIC or IV ovarian cancer. *N Engl J Med* 2010; **363**(10): 943-53.

87. Kehoe S, Hook J, Nankivell M, et al. Primary chemotherapy versus primary surgery for newly diagnosed advanced ovarian cancer (CHORUS): an open-label, randomised, controlled, non-inferiority trial. *Lancet* 2015; **386**(9990): 249-57.
88. Vergote I, Coens C, Nankivell M, et al. Neoadjuvant chemotherapy versus debulking surgery in advanced tubo-ovarian cancers: pooled analysis of individual patient data from the EORTC 55971 and CHORUS trials. *Lancet Oncol* 2018; **19**(12): 1680-7.
89. Marchesi JR, Ravel J. The vocabulary of microbiome research: a proposal. *Microbiome* 2015; **3**.
90. Human Microbiome Project C. Structure, function and diversity of the healthy human microbiome. *Nature* 2012; **486**(7402): 207-14.
91. Integrative HMPRNC. The Integrative Human Microbiome Project: dynamic analysis of microbiome-host omics profiles during periods of human health and disease. *Cell Host Microbe* 2014; **16**(3): 276-89.
92. Sender R, Fuchs S, Milo R. Are We Really Vastly Outnumbered? Revisiting the Ratio of Bacterial to Host Cells in Humans. *Cell* 2016; **164**(3): 337-40.
93. Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R, Gordon JI. The human microbiome project. *Nature* 2007; **449**.
94. Ottman N, Smidt H, de Vos WM, Belzer C. The function of our microbiota: who is out there and what do they do? *Front Cell Infect Microbiol* 2012; **2**: 104.
95. Gilbert JA, Blaser MJ, Caporaso JG, Jansson JK, Lynch SV, Knight R. Current understanding of the human microbiome. *Nat Med* 2018; **24**(4): 392-400.
96. Japanese Microbiome Consortium (JMBC). https://jmbc.life/en/index.html.
97. Canadian Microbiome Initiative 1/2. https://cihr-irsc.gc.ca/e/51499.html.
98. Blanco-Miguez A, Gutierrez-Jacome A, Fdez-Riverola F, Lourenco A, Sanchez B. MAHMI database: a comprehensive MetaHit-based resource for the study of the mechanism of action of the human microbiota. *Database (Oxford)* 2017; **2017**.
99. Cusack S, O'Toole PW, consortium E. Challenges and implications for biomedical research and intervention studies in older populations: insights from the ELDERMET study. *Gerontology* 2013; **59**(2): 114-21.
100. Ravel J, Gajer P, Abdo Z, et al. Vaginal microbiome of reproductive-age women. *Proc Natl Acad Sci U S A* 2011; **108 Suppl 1**: 4680-7.
101. Gajer P, Brotman RM, Bai G, et al. Temporal dynamics of the human vaginal microbiota. *Sci Transl Med* 2012; **4**.
102. Chen Y, Qiu X, Wang W, et al. Human papillomavirus infection and cervical intraepithelial neoplasia progression are associated with increased vaginal microbiome diversity in a Chinese cohort. *BMC Infect Dis* 2020; **20**(1): 629.
103. Hakimjavadi H, George SH, Taub M, et al. The vaginal microbiome is associated with endometrial cancer grade and histology. *Cancer Res Commun* 2022; **2**(6): 447-55.
104. Mitra A, MacIntyre DA, Lee YS, et al. Cervical intraepithelial neoplasia disease progression is associated with increased vaginal microbiome diversity. *Sci Rep* 2015; **5**.
105. 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**(1): 122.
106. 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**(1): 875.
107. Verstraelen H, Vilchez-Vargas R, Desimpel F, et al. Characterisation of the human uterine microbiome in non-pregnant women through deep sequencing of the V1-2 region of the 16S rRNA gene. *PeerJ* 2016; **4**: e1602.
108. Miles SM, Hardy BL, Merrell DS. Investigation of the microbiota of the reproductive tract in women undergoing a total hysterectomy and bilateral salpingo-oopherectomy. *Fertility and Sterility* 2017; **107**(3): 813-20.e1.
109. Mitchell CM, Haick A, Nkwopara E, et al. Colonization of the upper genital tract by vaginal bacterial species in nonpregnant women. *Am J Obstet Gynecol* 2015; **212**(5): 611 e1-9.
110. Walsh DM, Hokenstad AN, Chen J, et al. Postmenopause as a key factor in the composition of the Endometrial Cancer Microbiome (ECbiome). *Sci Rep* 2019; **9**(1): 19213.

111. Miles SM, Hardy BL, Merrell DS. Investigation of the microbiota of the reproductive tract in women undergoing a total hysterectomy and bilateral salpingo-oopherectomy. *Fertil Steril* 2017; **107**(3): 813-20 e1.
112. Schwabe RF, Jobin C. The microbiome and cancer. *Nat Rev Cancer* 2013; **13**(11): 800-12.
113. Garrett WS. Cancer and the microbiota. *Science* 2015; **348**(6230): 80-6.
114. Yang TK, Chung CJ, Chung SD, Muo CH, Chang CH, Huang CY. Risk of Endometrial Cancer in Women With Pelvic Inflammatory Disease: A Nationwide Population-Based Retrospective Cohort Study. *Medicine (Baltimore)* 2015; **94**(34): e1278.
115. Ness RB, Cottreau C. Possible role of ovarian epithelial inflammation in ovarian cancer. *J Natl Cancer Inst* 1999; **91**(17): 1459-67.
116. Maccio A, Madeddu C. Inflammation and ovarian cancer. *Cytokine* 2012; **58**(2): 133-47.
117. Cole AJ, Dwight T, Gill AJ, et al. Assessing mutant p53 in primary high-grade serous ovarian cancer using immunohistochemistry and massively parallel sequencing. *Sci Rep* 2016; **6**: 26191.
118. Uehara I, Tanaka N. Role of p53 in the Regulation of the Inflammatory Tumor Microenvironment and Tumor Suppression. *Cancers (Basel)* 2018; **10**(7).
119. Zhou Y, Chlebowski R, LaMonte MJ, et al. Body mass index, physical activity, and mortality in women diagnosed with ovarian cancer: results from the Women's Health Initiative. *Gynecol Oncol* 2014; **133**(1): 4-10.
120. Sterne JA DSG. Sifting the evidence. *Proc Natl Acad Sci U S A* 2001; **322**(7280): 226-31.
121. Johnson VE. Revised standards for statistical evidence. *Proc Natl Acad Sci U S A* 2013; **110**(48): 19313-7.
122. Ioannidis JP TR, McLaughlin JK. The false-positive to false-negative ratio in epidemiologic studies. *Epidemiology* 2011; **22**(4): 450-6.
123. Kalliala I, Markozannes G, Gunter MJ, et al. Obesity and gynaecological and obstetric conditions: umbrella review of the literature. *Bmj* 2017; **359**: j4511.
124. Kyrgiou M, Kalliala I, Markozannes G, et al. Adiposity and cancer at major anatomical sites: umbrella review of the literature. *BMJ* 2017; **356**: j477.
125. Raglan O, Kalliala I, Markozannes G, et al. Risk factors for endometrial cancer: An umbrella review of the literature. *Int J Cancer* 2019; **145**(7): 1719-30.
126. Shea BJ, Reeves BC, Wells G, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ* 2017; **358**: j4008.
127. DerSimonian R LN. Meta-analysis in clinical trials. . *Controlled clinical trials* 1986; **7**(3): 177-88.
128. WG C. The combination of estimates from different experiments. *Biometrics* 1954; **10**: 101-29.
129. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003; **327**(7414): 557-60.
130. Ioannidis JP, Trikalinos TA. An exploratory test for an excess of significant findings. *Clin Trials* 2007; **4**(3): 245-53.
131. IntHout J, Ioannidis JP, Rovers MM, Goeman JJ. Plea for routinely presenting prediction intervals in meta-analysis. *BMJ Open* 2016; **6**(7): e010247.
132. Riley RD, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. *BMJ* 2011; **342**: d549.
133. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997; **315**(7109): 629-34.
134. Nuesch E, Trelle S, Reichenbach S, et al. Small study effects in meta-analyses of osteoarthritis trials: meta-epidemiological study. *BMJ* 2010; **341**: c3515.
135. Tsilidis KK, Papatheodorou SI, Evangelou E, Ioannidis JP. Evaluation of excess statistical significance in meta-analyses of 98 biomarker associations with cancer risk. *J Natl Cancer Inst* 2012; **104**(24): 1867-78.
136. Salanti G, Ioannidis JP. Synthesis of observational studies should consider credibility ceilings. *J Clin Epidemiol* 2009; **62**(2): 115-22.

137. Papatheodorou SI, Tsilidis KK, Evangelou E, Ioannidis JP. Application of credibility ceilings probes the robustness of meta-analyses of biomarkers and cancer risk. *J Clin Epidemiol* 2015; **68**(2): 163-74.
138. Corporation S. Statistical Software: Release 15.0. College Station, Texas; 2015.
139. MacIntyre DA, Chandiramani M, Lee YS, et al. The vaginal microbiome during pregnancy and the postpartum period in a European population. *Sci Rep* 2015; **5**.
140. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl Environ Microbiol* 2013; **79**(17): 5112-20.
141. Wang Q, Garrity GM, Tiedje JM, Cole JR. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* 2007; **73**(16): 5261-7.
142. Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 2010; **26**(19): 2460-1.
143. Winters AD, Romero R, Gervasi MT, et al. Does the endometrial cavity have a molecular microbial signature? *Sci Rep* 2019; **9**(1): 9905.
144. Parks DH, Beiko RG. Identifying biologically relevant differences between metagenomic communities. *Bioinformatics* 2010; **26**(6): 715-21.
145. Dhariwal A, Chong J, Habib S, King IL, Agellon LB, Xia J. MicrobiomeAnalyst: a web-based tool for comprehensive statistical, visual and meta-analysis of microbiome data. *Nucleic Acids Res* 2017; **45**(W1): W180-W8.
146. Chong J, Liu P, Zhou G, Xia J. Using MicrobiomeAnalyst for comprehensive statistical, functional, and meta-analysis of microbiome data. *Nat Protoc* 2020; **15**(3): 799-821.
147. France MT, Ma B, Gajer P, et al. VALENCIA: a nearest centroid classification method for vaginal microbial communities based on composition. *Microbiome* 2020; **8**(1): 166.
148. Davis NM, Proctor DM, Holmes SP, Relman DA, Callahan BJ. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome* 2018; **6**(1): 226.
149. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018; **68**(6): 394-424.
150. Ferlay J. EM, Lam F., Colombet M., Mery L., Piñeros M., Znaor A., Soerjomataram I., Bray F. Global Cancer Observatory: Cancer Tomorrow. 2018. https://gco.iarc.fr/tomorrow (accessed 16 March 2020 2020).
151. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *CA: A Cancer Journal for Clinicians* 2018; **68**(1): 7-30.
152. Menon U, Gentry-Maharaj A, Hallett R, et al. Sensitivity and specificity of multimodal and ultrasound screening for ovarian cancer, and stage distribution of detected cancers: results of the prevalence screen of the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS). *Lancet Oncol* 2009; **10**(4): 327-40.
153. Collaborative Group On Epidemiological Studies Of Ovarian C, Beral V, Gaitskell K, et al. Menopausal hormone use and ovarian cancer risk: individual participant meta-analysis of 52 epidemiological studies. *Lancet* 2015; **385**(9980): 1835-42.
154. Ioannidis JPA. Why Most Published Research Findings Are False. *PLOS Medicine* 2005; **2**(8): e124.
155. Ioannidis JPA. Why Most Discovered True Associations Are Inflated. *Epidemiology* 2008; **19**(5): 640-8.
156. Dwan K, Gamble C, Williamson PR, Kirkham JJ. Systematic Review of the Empirical Evidence of Study Publication Bias and Outcome Reporting Bias — An Updated Review. *PLOS ONE* 2013; **8**(7): e66844.
157. Theodoratou E, Tzoulaki I, Zgaga L, Ioannidis JP. Vitamin D and multiple health outcomes: umbrella review of systematic reviews and meta-analyses of observational studies and randomised trials. *BMJ* 2014; **348**: g2035.

158. Belbasis L, Bellou V, Evangelou E, Ioannidis JP, Tzoulaki I. Environmental risk factors and multiple sclerosis: an umbrella review of systematic reviews and meta-analyses. *Lancet Neurol* 2015; **14**(3): 263-73.
159. Tsilidis KK, Kasimis JC, Lopez DS, Ntzani EE, Ioannidis JP. Type 2 diabetes and cancer: umbrella review of meta-analyses of observational studies. *BMJ* 2015; **350**: g7607.
160. Rezende LFM, Sa TH, Markozannes G, et al. Physical activity and cancer: an umbrella review of the literature including 22 major anatomical sites and 770 000 cancer cases. *Br J Sports Med* 2018; **52**(13): 826-33.
161. Aune D, Navarro Rosenblatt DA, Chan DS, et al. Anthropometric factors and ovarian cancer risk: a systematic review and nonlinear dose-response meta-analysis of prospective studies. *Int J Cancer* 2015; **136**(8): 1888-98.
162. Baandrup L, Faber MT, Christensen J, et al. Nonsteroidal anti-inflammatory drugs and risk of ovarian cancer: systematic review and meta-analysis of observational studies. *Acta Obstet Gynecol Scand* 2013; **92**(3): 245-55.
163. Barry JA, Azizia MM, Hardiman PJ. Risk of endometrial, ovarian and breast cancer in women with polycystic ovary syndrome: a systematic review and meta-analysis. *Hum Reprod Update* 2014; **20**(5): 748-58.
164. Berge W, Mundt K, Luu H, Boffetta P. Genital use of talc and risk of ovarian cancer: a meta-analysis. *Eur J Cancer Prev* 2017.
165. Bernatsky S, Ramsey-Goldman R, Foulkes WD, Gordon C, Clarke AE. Breast, ovarian, and endometrial malignancies in systemic lupus erythematosus: a meta-analysis. *Br J Cancer* 2011; **104**(9): 1478-81.
166. Bonovas S, Filioussi K, Sitaras NM. Paracetamol use and risk of ovarian cancer: a meta-analysis. *Br J Clin Pharmacol* 2006; **62**(1): 113-21.
167. Butler LM, Wu AH. Green and black tea in relation to gynecologic cancers. *Mol Nutr Food Res* 2011; **55**(6): 931-40.
168. Camargo MC, Stayner LT, Straif K, et al. Occupational exposure to asbestos and ovarian cancer: a meta-analysis. *Environ Health Perspect* 2011; **119**(9): 1211-7.
169. Collaborative Group on Epidemiological Studies of Ovarian C, 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**(9609): 303-14.
170. Gao M, Ma W, Chen XB, Chang ZW, Zhang XD, Zhang MZ. Meta-analysis of green tea drinking and the prevalence of gynecological tumors in women. *Asia Pac J Public Health* 2013; **25**(4 Suppl): 43S-8S.
171. Garg PP, Kerlikowske K, Subak L, Grady D. Hormone replacement therapy and the risk of epithelial ovarian carcinoma: a meta-analysis. *Obstet Gynecol* 1998; **92**(3): 472-9.
172. Gong TT, Wu QJ, Vogtmann E, Lin B, Wang YL. Age at menarche and risk of ovarian cancer: a meta-analysis of epidemiological studies. *Int J Cancer* 2013; **132**(12): 2894-900.
173. Han B, Li X, Yu T. Cruciferous vegetables consumption and the risk of ovarian cancer: a meta-analysis of observational studies. *Diagn Pathol* 2014; **9**: 7.
174. Hankinson SE, Colditz GA, Hunter DJ, Spencer TL, Rosner B, Stampfer MJ. A quantitative assessment of oral contraceptive use and risk of ovarian cancer. *Obstet Gynecol* 1992; **80**(4): 708-14.
175. Huncharek M, Klassen H, Kupelnick B. Dietary beta-carotene intake and the risk of epithelial ovarian cancer: a meta-analysis of 3,782 subjects from five observational studies. *In Vivo* 2001; **15**(4): 339-43.
176. Huncharek M, Muscat J, Onitilo A, Kupelnick B. Use of cosmetic talc on contraceptive diaphragms and risk of ovarian cancer: a meta-analysis of nine observational studies. *Eur J Cancer Prev* 2007; **16**(5): 422-9.
177. Huo YL, Qiao JM, Gao S. Association between antidepressant medication use and epithelial ovarian cancer risk: a systematic review and meta-analysis of observational studies. *Br J Clin Pharmacol* 2018; **84**(4): 649-58.

178. Ip S, Chung M, Raman G, et al. Breastfeeding and maternal and infant health outcomes in developed countries. *Evidence Report/Technology Assessment* 2007; (153): 1-186.
179. Jiang PY, Jiang ZB, Shen KX, Yue Y. Fish intake and ovarian cancer risk: a meta-analysis of 15 case-control and cohort studies. *PLoS One* 2014; **9**(4): e94601.
180. Keum N, Greenwood DC, Lee DH, et al. Adult weight gain and adiposity-related cancers: a dose-response meta-analysis of prospective observational studies. *J Natl Cancer Inst* 2015; **107**(2).
181. Keum N, Lee DH, Marchand N, et al. Egg intake and cancers of the breast, ovary and prostate: a dose-response meta-analysis of prospective observational studies. *Br J Nutr* 2015; **114**(7): 1099-107.
182. Kim HS, Kim JW, Shouten LJ, et al. Wine drinking and epithelial ovarian cancer risk: a meta-analysis. *J Gynecol Oncol* 2010; **21**(2): 112-8.
183. Kim HS, Kim TH, Chung HH, Song YS. Risk and prognosis of ovarian cancer in women with endometriosis: a meta-analysis. *Br J Cancer* 2014; **110**(7): 1878-90.
184. Kolahdooz F, van der Pols JC, Bain CJ, et al. Meat, fish, and ovarian cancer risk: Results from 2 Australian case-control studies, a systematic review, and meta-analysis. *Am J Clin Nutr* 2010; **91**(6): 1752-63.
185. Koushik A, Hunter DJ, Spiegelman D, et al. Intake of the major carotenoids and the risk of epithelial ovarian cancer in a pooled analysis of 10 cohort studies. *Int J Cancer* 2006; **119**(9): 2148-54.
186. Larsson SC, Orsini N, Wolk A. Milk, milk products and lactose intake and ovarian cancer risk: a meta-analysis of epidemiological studies. *Int J Cancer* 2006; **118**(2): 431-41.
187. Lee PN, Thornton AJ, Hamling JS. Epidemiological evidence on environmental tobacco smoke and cancers other than lung or breast. *Regulatory Toxicology & Pharmacology* 2016; **80**: 134-63.
188. Li DP, Du C, Zhang ZM, et al. Breastfeeding and ovarian cancer risk: a systematic review and meta-analysis of 40 epidemiological studies. *Asian Pac J Cancer Prev* 2014; **15**(12): 4829-37.
189. Li X, Xu J. Meta-analysis of the association between dietary lycopene intake and ovarian cancer risk in postmenopausal women. *Sci Rep* 2014; **4**: 4885.
190. Li J, Jiao X, Yuan Z, Qiu H, Guo R. C-reactive protein and risk of ovarian cancer: A systematic review and meta-analysis. *Medicine (Baltimore)* 2017; **96**(34): e7822.
191. Liu Y, Qin A, Li T, Qin X, Li S. Effect of statin on risk of gynecologic cancers: a meta-analysis of observational studies and randomized controlled trials. *Gynecol Oncol* 2014; **133**(3): 647-55.
192. Liu Z, Zhang TT, Zhao JJ, et al. The association between overweight, obesity and ovarian cancer: a meta-analysis. *Jpn J Clin Oncol* 2015; **45**(12): 1107-15.
193. Liu ZY, Gao XP, Zhu S, et al. Dietary inflammatory index and risk of gynecological cancers: a systematic review and meta-analysis of observational studies. *J Gynecol Oncol* 2019; **30**(3): e23.
194. Luan NN, Wu QJ, Gong TT, Vogtmann E, Wang YL, Lin B. Breastfeeding and ovarian cancer risk: a meta-analysis of epidemiologic studies. *Am J Clin Nutr* 2013; **98**(4): 1020-31.
195. Murphy MA, Trabert B, Yang HP, et al. Non-steroidal anti-inflammatory drug use and ovarian cancer risk: findings from the NIH-AARP Diet and Health Study and systematic review. *Cancer Causes Control* 2012; **23**(11): 1839-52.
196. Negri E, Franceschi S, Tzonou A, et al. Pooled analysis of 3 European case-control studies: I. Reproductive factors and risk of epithelial ovarian cancer. *Int J Cancer* 1991; **49**(1): 50-6.
197. Olsen CM, Green AC, Whiteman DC, Sadeghi S, Kolahdooz F, Webb PM. Obesity and the risk of epithelial ovarian cancer: a systematic review and meta-analysis. *Eur J Cancer* 2007; **43**(4): 690-709.
198. Pang Y, Wang W. Dietary protein intake and risk of ovarian cancer: evidence from a meta-analysis of observational studies. *Biosci Rep* 2018; **38**(6).
199. Park B, Park S, Shin HR, et al. Population attributable risks of modifiable reproductive factors for breast and ovarian cancers in Korea. *BMC Cancer* 2016; **16**: 5.
200. Pelucchi C, Bosetti C, Galeone C, La Vecchia C. Dietary acrylamide and cancer risk: an updated meta-analysis. *Int J Cancer* 2015; **136**(12): 2912-22.
201. Penninkilampi R, Eslick GD. Perineal Talc Use and Ovarian Cancer: A Systematic Review and Meta-Analysis. *Epidemiology* 2018; **29**(1): 41-9.
202. Poorolajal J, Jenabi E, Masoumi SZ. Body mass index effects on risk of ovarian cancer: a meta-analysis. *Asian Pac J Cancer Prev* 2014; **15**(18): 7665-71.

203. Qiu W, Lu H, Qi Y, Wang X. Dietary fat intake and ovarian cancer risk: a meta-analysis of epidemiological studies. *Oncotarget* 2016; **7**(24): 37390-406.
204. Qu XL, Fang Y, Zhang M, Zhang YZ. Phytoestrogen intake and risk of ovarian cancer: a meta-analysis of 10 observational studies. *Asian Pac J Cancer Prev* 2014; **15**(21): 9085-91.
205. Shafiei F, Salari-Moghaddam A, Milajerdi A, Larijani B, Esmaillzadeh A. Coffee and caffeine intake and risk of ovarian cancer: a systematic review and meta-analysis. *Int J Gynecol Cancer* 2019; **29**(3): 579-84.
206. Schmid D, Leitzmann MF. Television viewing and time spent sedentary in relation to cancer risk: a meta-analysis. *J Natl Cancer Inst* 2014; **106**(7).
207. Shi LF, Wu Y, Li CY. Hormone therapy and risk of ovarian cancer in postmenopausal women: a systematic review and meta-analysis. *Menopause* 2016; **23**(4): 417-24.
208. Siristatidis C, Sergentanis TN, Kanavidis P, et al. Controlled ovarian hyperstimulation for IVF: impact on ovarian, endometrial and cervical cancer--a systematic review and meta-analysis. *Hum Reprod Update* 2013; **19**(2): 105-23.
209. Song X, Li Z, Ji X, Zhang D. Calcium Intake and the Risk of Ovarian Cancer: A Meta-Analysis. *Nutrients* 2017; **9**(7).
210. Song L, Wang Y, Zhang J, Song N, Xu X, Lu Y. The risks of cancer development in systemic lupus erythematosus (SLE) patients: a systematic review and meta-analysis. *Arthritis Res Ther* 2018; **20**(1): 270.
211. Wallin A, Orsini N, Wolk A. Red and processed meat consumption and risk of ovarian cancer: a dose-response meta-analysis of prospective studies. *Br J Cancer* 2011; **104**(7): 1196-201.
212. Wang HF, Yao AL, Sun YY, Zhang AH. Empirically derived dietary patterns and ovarian cancer risk: a meta- analysis. *Eur J Cancer Prev* 2018; **27**(5): 493-501.
213. Wen Q, Zhao Z, Wen J, et al. The association between metformin therapy and risk of gynecological cancer in patients: Two meta-analyses. *Eur J Obstet Gynecol Reprod Biol* 2019; **237**: 33-41.
214. Xu H, Ding Y, Xin X, Wang W, Zhang D. Dietary fiber intake is associated with a reduced risk of ovarian cancer: a dose-response meta-analysis. *Nutr Res* 2018; **57**: 1-11.
215. Yan-Hong H, Jing L, Hong L, Shan-Shan H, Yan L, Ju L. Association between alcohol consumption and the risk of ovarian cancer: a meta-analysis of prospective observational studies. *BMC Public Health* 2015; **15**: 223.
216. Yin L, Grandi N, Raum E, Haug U, Arndt V, Brenner H. Meta-analysis: Circulating vitamin D and ovarian cancer risk. *Gynecol Oncol* 2011; **121**(2): 369-75.
217. Zeng ST, Guo L, Liu SK, et al. Egg consumption is associated with increased risk of ovarian cancer: Evidence from a meta-analysis of observational studies. *Clin Nutr* 2015; **34**(4): 635-41.
218. Zeng F, Wei H, Yeoh E, et al. Inflammatory Markers of CRP, IL6, TNFalpha, and Soluble TNFR2 and the Risk of Ovarian Cancer: A Meta-analysis of Prospective Studies. *Cancer Epidemiol Biomarkers Prev* 2016; **25**(8): 1231-9.
219. Zhan X, Wang J, Pan S, Lu C. Tea consumption and the risk of ovarian cancer: A meta-analysis of epidemiological studies. *Oncotarget* 2017.
220. Zhang D, Bai B, Xi Y, Wang T, Zhao Y. Is aspirin use associated with a decreased risk of ovarian cancer? A systematic review and meta-analysis of observational studies with dose-response analysis. *Gynecol Oncol* 2016; **142**(2): 368-77.
221. Zhang D, Li N, Xi Y, Zhao Y, Wang T. Diabetes mellitus and risk of ovarian cancer. A systematic review and meta-analysis of 15 cohort studies. *Diabetes Res Clin Pract* 2017; **130**: 43-52.
222. Zhang D, Kaushiva A, Xi Y, Wang T, Li N. Non-herbal tea consumption and ovarian cancer risk: a systematic review and meta-analysis of observational epidemiologic studies with indirect comparison and dose-response analysis. *Carcinogenesis* 2018; **39**(6): 808-18.
223. Zhang XS, Zhang YM, Li B, Fan B, Zhao Y, Yang SJ. Risk reduction of endometrial and ovarian cancer after bisphosphonates use: A meta-analysis. *Gynecol Oncol* 2018; **150**(3): 509-14.
224. Zheng B, Shen H, Han H, Han T, Qin Y. Dietary fiber intake and reduced risk of ovarian cancer: a meta-analysis. *Nutr J* 2018; **17**(1): 99.
225. Zhong S, Chen L, Lv M, Ma T, Zhang X, Zhao J. Nonoccupational physical activity and risk of ovarian cancer: a meta-analysis. *Tumour Biol* 2014; **35**(11): 11065-73.

226. Zhou B, Sun Q, Cong R, et al. Hormone replacement therapy and ovarian cancer risk: a meta-analysis. *Gynecol Oncol* 2008; **108**(3): 641-51.
227. Wu Y, Sun W, Xin X, Wang W, Zhang D. Age at last birth and risk of developing epithelial ovarian cancer: a meta-analysis. *Biosci Rep* 2019; **39**(9).
228. Salari-Moghaddam A, Milajerdi A, Surkan PJ, Larijani B, Esmaillzadeh A. Caffeine, Type of Coffee, and Risk of Ovarian Cancer: A Dose-Response Meta-Analysis of Prospective Studies. *J Clin Endocrinol Metab* 2019; **104**(11): 5349-59.
229. Santucci C, Bosetti C, Peveri G, et al. Dose-risk relationships between cigarette smoking and ovarian cancer histotypes: a comprehensive meta-analysis. *Cancer Causes Control* 2019; **30**(9): 1023-32.
230. He Y, Li X, Gasevic D, et al. Statins and Multiple Noncardiovascular Outcomes: Umbrella Review of Meta-analyses of Observational Studies and Randomized Controlled Trials. *Ann Intern Med* 2018; **169**(8): 543-53.
231. Li J, Liu R, Tang S, et al. Impact of endometriosis on risk of ovarian, endometrial and cervical cancers: a meta-analysis. *Arch Gynecol Obstet* 2019; **299**(1): 35-46.
232. Ioannidis JP. Integration of evidence from multiple meta-analyses: a primer on umbrella reviews, treatment networks and multiple treatments meta-analyses. *CMAJ* 2009; **181**(8): 488-93.
233. Giovannucci E. A growing link-what is the role of height in cancer risk? *Br J Cancer* 2019; **120**(6): 575-6.
234. Pollak MN, Schernhammer ES, Hankinson SE. Insulin-like growth factors and neoplasia. *Nat Rev Cancer* 2004; **4**(7): 505-18.
235. Yarmolinsky J, Relton CL, Lophatananon A, et al. Appraising the role of previously reported risk factors in epithelial ovarian cancer risk: A Mendelian randomization analysis. *PLoS Med* 2019; **16**(8): e1002893.
236. Olsen CM, Nagle CM, Whiteman DC, et al. Obesity and risk of ovarian cancer subtypes: evidence from the Ovarian Cancer Association Consortium. *Endocr Relat Cancer* 2013; **20**(2): 251-62.
237. Lauby-Secretan B, Scoccianti C, Loomis D, et al. Body Fatness and Cancer--Viewpoint of the IARC Working Group. *N Engl J Med* 2016; **375**(8): 794-8.
238. Avgerinos KI, Spyrou N, Mantzoros CS, Dalamaga M. Obesity and cancer risk: Emerging biological mechanisms and perspectives. *Metabolism* 2019; **92**: 121-35.
239. Charlton BM, Rich-Edwards JW, Colditz GA, et al. Oral contraceptive use and mortality after 36 years of follow-up in the Nurses' Health Study: prospective cohort study. *BMJ* 2014; **349**: g6356.
240. Iversen L, Sivasubramaniam S, Lee AJ, Fielding S, Hannaford PC. Lifetime cancer risk and combined oral contraceptives: the Royal College of General Practitioners' Oral Contraception Study. *Am J Obstet Gynecol* 2017; **216**(6): 580 e1- e9.
241. McGuire V, Hartge P, Liao LM, et al. Parity and Oral Contraceptive Use in Relation to Ovarian Cancer Risk in Older Women. *Cancer Epidemiol Biomarkers Prev* 2016; **25**(7): 1059-63.
242. Michels KA, Trabert B. Oral Contraceptive Progestin and Estrogen Use and Increases in Breast, Ovarian, and Endometrial Cancers-Reply. *JAMA Oncol* 2018; **4**(11): 1623-4.
243. Gaitskell K, Green J, Pirie K, et al. Histological subtypes of ovarian cancer associated with parity and breastfeeding in the prospective Million Women Study. *Int J Cancer* 2018; **142**(2): 281-9.
244. Cunat S, Hoffmann P, Pujol P. Estrogens and epithelial ovarian cancer. *Gynecol Oncol* 2004; **94**(1): 25-32.
245. Mungenast F, Thalhammer T. Estrogen biosynthesis and action in ovarian cancer. *Front Endocrinol (Lausanne)* 2014; **5**: 192.
246. Lau KM, Mok SC, Ho SM. Expression of human estrogen receptor-alpha and -beta, progesterone receptor, and androgen receptor mRNA in normal and malignant ovarian epithelial cells. *Proc Natl Acad Sci U S A* 1999; **96**(10): 5722-7.
247. Tsilidis KK, Allen NE, Key TJ, et al. Menopausal hormone therapy and risk of ovarian cancer in the European prospective investigation into cancer and nutrition. *Cancer Causes Control* 2011; **22**(8): 1075-84.
248. Hildebrand JS, Gapstur SM, Feigelson HS, Teras LR, Thun MJ, Patel AV. Postmenopausal hormone use and incident ovarian cancer: Associations differ by regimen. *Int J Cancer* 2010; **127**(12): 2928-35.

249. Iversen L, Fielding S, Lidegaard O, Morch LS, Skovlund CW, Hannaford PC. Association between contemporary hormonal contraception and ovarian cancer in women of reproductive age in Denmark: prospective, nationwide cohort study. *BMJ* 2018; **362**: k3609.
250. Vessey M, Yeates D. Oral contraceptive use and cancer: final report from the Oxford-Family Planning Association contraceptive study. *Contraception* 2013; **88**(6): 678-83.
251. Tsilidis KK, Allen NE, Key TJ, et al. Oral contraceptive use and reproductive factors and risk of ovarian cancer in the European Prospective Investigation into Cancer and Nutrition. *Br J Cancer* 2011; **105**(9): 1436-42.
252. Pearce CL, Templeman C, Rossing MA, et al. Association between endometriosis and risk of histological subtypes of ovarian cancer: a pooled analysis of case-control studies. *Lancet Oncol* 2012; **13**(4): 385-94.
253. Pavone ME, Lyttle BM. Endometriosis and ovarian cancer: links, risks, and challenges faced. *Int J Womens Health* 2015; **7**: 663-72.
254. Dawson A, Fernandez ML, Anglesio M, Yong PJ, Carey MS. Endometriosis and endometriosis-associated cancers: new insights into the molecular mechanisms of ovarian cancer development. *Ecancermedicalscience* 2018; **12**: 803.
255. Zhang Z, Zhai Z, Yang Y, et al. Diabetes mellitus is associated with increased bleeding in pulmonary embolism receiving conventional anticoagulant therapy: findings from a "real-world" study. *J Thromb Thrombolysis* 2017; **43**(4): 540-9.
256. Narod SA. Talc and ovarian cancer. *Gynecol Oncol* 2016; **141**(3): 410-2.
257. O'Brien KM, Tworoger SS, Harris HR, et al. Association of Powder Use in the Genital Area With Risk of Ovarian Cancer. *JAMA* 2020; **323**(1): 49-59.
258. Segata N, Haake SK, Mannon P, et al. Composition of the adult digestive tract bacterial microbiome based on seven mouth surfaces, tonsils, throat and stool samples. *Genome Biol* 2012; **13**(6): R42.
259. Eckburg PB, Bik EM, Bernstein CN, et al. Diversity of the human intestinal microbial flora. *Science* 2005; **308**(5728): 1635-8.
260. Grice EA, Kong HH, Conlan S, et al. Topographical and temporal diversity of the human skin microbiome. *Science* 2009; **324**(5931): 1190-2.
261. Dewhirst FE, Chen T, Izard J, et al. The human oral microbiome. *J Bacteriol* 2010; **192**(19): 5002-17.
262. Costello EK, Lauber CL, Hamady M, Fierer N, Gordon JI, Knight R. Bacterial community variation in human body habitats across space and time. *Science* 2009; **326**(5960): 1694-7.
263. Romero R, Hassan SS, Gajer P, et al. The composition and stability of the vaginal microbiota of normal pregnant women is different from that of non-pregnant women. *Microbiome* 2014; **2**.
264. Audirac-Chalifour A, Torres-Poveda K, Bahena-Roman M, et al. Cervical microbiome and cytokine profile at various stages of cervical cancer: a pilot study. *PLoS One* 2016; **11**.
265. de Goffau MC, Lager S, Salter SJ, et al. Recognizing the reagent microbiome. *Nat Microbiol* 2018; **3**(8): 851-3.
266. Jervis-Bardy J, Leong LE, Marri S, et al. Deriving accurate microbiota profiles from human samples with low bacterial content through post-sequencing processing of Illumina MiSeq data. *Microbiome* 2015; **3**: 19.
267. Moller BR, Kristiansen FV, Thorsen P, Frost L, Mogensen SC. Sterility of the uterine cavity. *Acta Obstet Gynecol Scand* 1995; **74**(3): 216-9.
268. Moreno I, Codoñer FM, Vilella F, et al. Evidence that the endometrial microbiota has an effect on implantation success or failure. *American Journal of Obstetrics and Gynecology* 2016; **215**(6): 684-703.
269. Kyono K, Hashimoto T, Nagai Y, Sakuraba Y. Analysis of endometrial microbiota by 16S ribosomal RNA gene sequencing among infertile patients: a single-center pilot study. *Reprod Med Biol* 2018; **17**(3): 297-306.
270. de Goffau MC, Lager S, Sovio U, et al. Human placenta has no microbiome but can contain potential pathogens. *Nature* 2019; **572**(7769): 329-34.

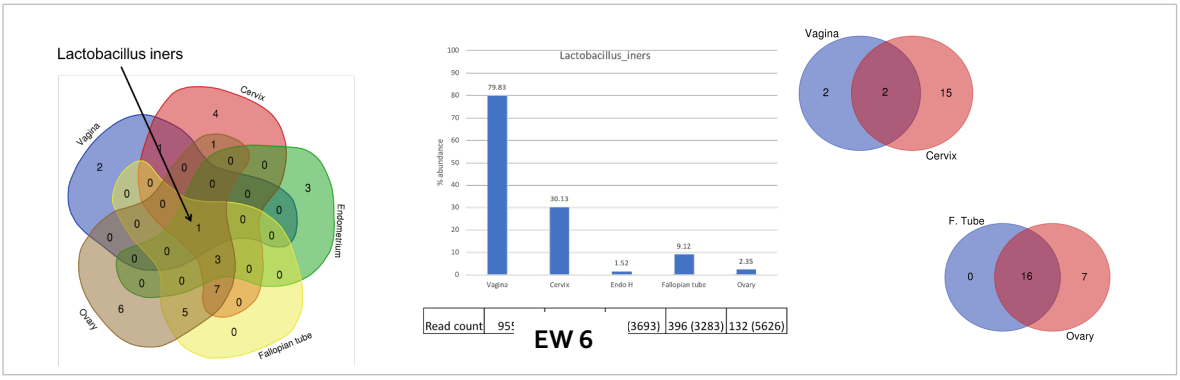
271. Erridge C, Moncayo-Nieto OL, Morgan R, Young M, Poxton IR. Acinetobacter baumannii lipopolysaccharides are potent stimulators of human monocyte activation via Toll-like receptor 4 signalling. *J Med Microbiol* 2007; **56**(Pt 2): 165-71.
272. Laniewski P, Ilhan ZE, Herbst-Kralovetz MM. The microbiome and gynaecological cancer development, prevention and therapy. *Nat Rev Urol* 2020; **17**(4): 232-50.
273. Jeon SJ, Cunha F, Vieira-Neto A, et al. Blood as a route of transmission of uterine pathogens from the gut to the uterus in cows. *Microbiome* 2017; **5**(1): 109.
274. Fardini Y, Chung P, Dumm R, Joshi N, Han YW. Transmission of diverse oral bacteria to murine placenta: evidence for the oral microbiome as a potential source of intrauterine infection. *Infect Immun* 2010; **78**(4): 1789-96.
275. Peric A, Weiss J, Vulliemoz N, Baud D, Stojanov M. Bacterial Colonization of the Female Upper Genital Tract. *Int J Mol Sci* 2019; **20**(14).
276. Halfvarson J, Brislawn CJ, Lamendella R, et al. Dynamics of the human gut microbiome in inflammatory bowel disease. *Nat Microbiol* 2017; **2**: 17004.
277. Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, Gordon JI. An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* 2006; **444**(7122): 1027-31.
278. Gajer P, Brotman RM, Bai G, et al. Temporal dynamics of the human vaginal microbiota. *Sci Transl Med* 2012; **4**(132): 132ra52.
279. Chandiramani M, Bennett PR, Brown R, Lee YS, Macintyre DA. Vaginal microbiome—pregnant host interactions determine a significant proportion of preterm labour. *Fetal Matern Med Rev* 2014; **25**.
280. Al-Memar M, Bobdiwala S, Fourie H, et al. The association between vaginal bacterial composition and miscarriage: a nested case-control study. *BJOG* 2020; **127**(2): 264-74.
281. Ravel J, Moreno I, Simon C. Bacterial vaginosis and its association with infertility, endometritis, and pelvic inflammatory disease. *Am J Obstet Gynecol* 2021; **224**(3): 251-7.
282. Mitra A, MacIntyre DA, Ntritsos G, et al. The vaginal microbiota associates with the regression of untreated cervical intraepithelial neoplasia 2 lesions. *Nat Commun* 2020; **11**(1): 1999.
283. Lu W, He F, Lin Z, et al. Dysbiosis of the endometrial microbiota and its association with inflammatory cytokines in endometrial cancer. *Int J Cancer* 2021; **148**(7): 1708-16.
284. Swidsinski A, Verstraelen H, Loening-Baucke V, Swidsinski S, Mendling W, Halwani Z. Presence of a polymicrobial endometrial biofilm in patients with bacterial vaginosis. *PLoS One* 2013; **8**(1): e53997.
285. Liu X, Mao B, Gu J, et al. Blautia-a new functional genus with potential probiotic properties? *Gut Microbes* 2021; **13**(1): 1-21.
286. Jumas-Bilak E, Jean-Pierre H, Carlier JP, et al. Dialister micraerophilus sp. nov. and Dialister propionicifaciens sp. nov., isolated from human clinical samples. *Int J Syst Evol Microbiol* 2005; **55**(Pt 6): 2471-8.
287. Bolton KL, Tyrer J, Song H, et al. Common variants at 19p13 are associated with susceptibility to ovarian cancer. *Nat Genet* 2010; **42**(10): 880-4.
288. Kuhn E, Kurman RJ, Vang R, et al. TP53 mutations in serous tubal intraepithelial carcinoma and concurrent pelvic high-grade serous carcinoma--evidence supporting the clonal relationship of the two lesions. *J Pathol* 2012; **226**(3): 421-6.
289. Kuhn E, Wang TL, Doberstein K, et al. CCNE1 amplification and centrosome number abnormality in serous tubal intraepithelial carcinoma: further evidence supporting its role as a precursor of ovarian high-grade serous carcinoma. *Mod Pathol* 2016; **29**(10): 1254-61.
290. Pradere JP, Dapito DH, Schwabe RF. The Yin and Yang of Toll-like receptors in cancer. *Oncogene* 2014; **33**(27): 3485-95.
291. Seidman JD, Sherman ME, Bell KA, Katabuchi H, O'Leary TJ, Kurman RJ. Salpingitis, salpingoliths, and serous tumors of the ovaries: is there a connection? *Int J Gynecol Pathol* 2002; **21**(2): 101-7.
292. Mandal S, Bandyopadhyay S, Tyagi K, Roy A. Human microbial dysbiosis as driver of gynecological malignancies. *Biochimie* 2022; **197**: 86-95.

293. Atashili J, Poole C, Ndumbe PM, Adimora AA, Smith JS. Bacterial vaginosis and HIV acquisition: a meta-analysis of published studies. *AIDS* 2008; **22**.
294. Gillet E, Meys JF, Verstraelen H, et al. Association between bacterial vaginosis and cervical intraepithelial neoplasia: systematic review and meta-analysis. *PLoS One* 2012; **7**.
295. Guo Y, You K, Qiao J, Zhao Y, Geng L. Bacterial vaginosis is conducive to the persistence of HPV infection. *Int J STD AIDS* 2012; **23**.
296. Taylor BD, Darville T, Haggerty CL. Does bacterial vaginosis cause pelvic inflammatory disease? *Sex Transm Dis* 2013; **40**(2): 117-22.
297. O'Callaghan JL, Turner R, Dekker Nitert M, Barrett HL, Clifton V, Pelzer ES. Re-assessing microbiomes in the low-biomass reproductive niche. *BJOG* 2020; **127**(2): 147-58.

APPENDIX

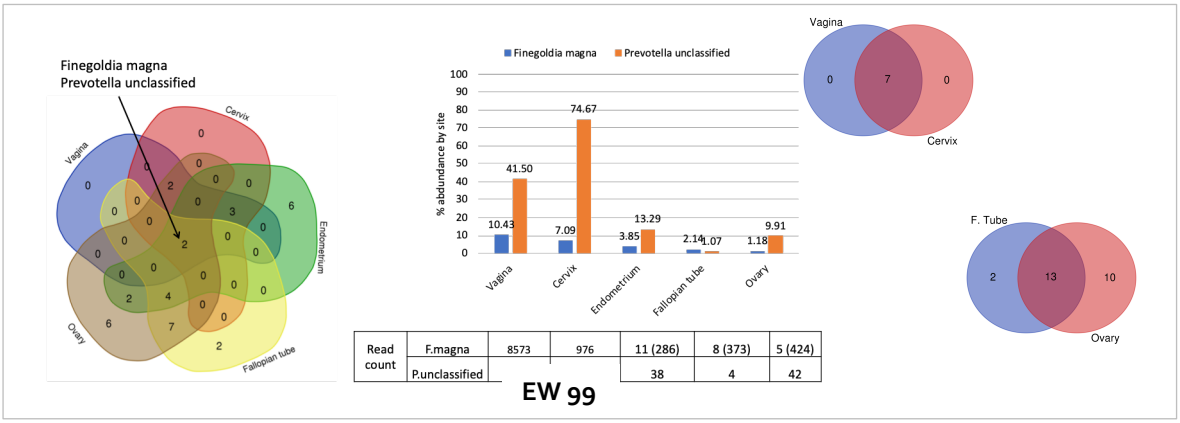
SUPPLEMENTARY RESULTS

Chapter 4- Common microbes in the genital tract; ***Malignant samples***



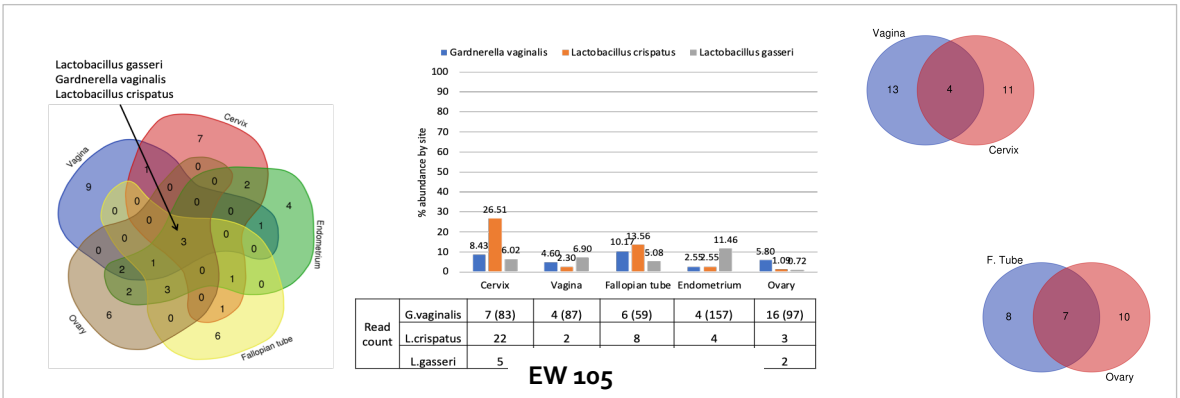
Read count	95[illegible]			[3693]	396 (3283)	132 (5626)

EW 6



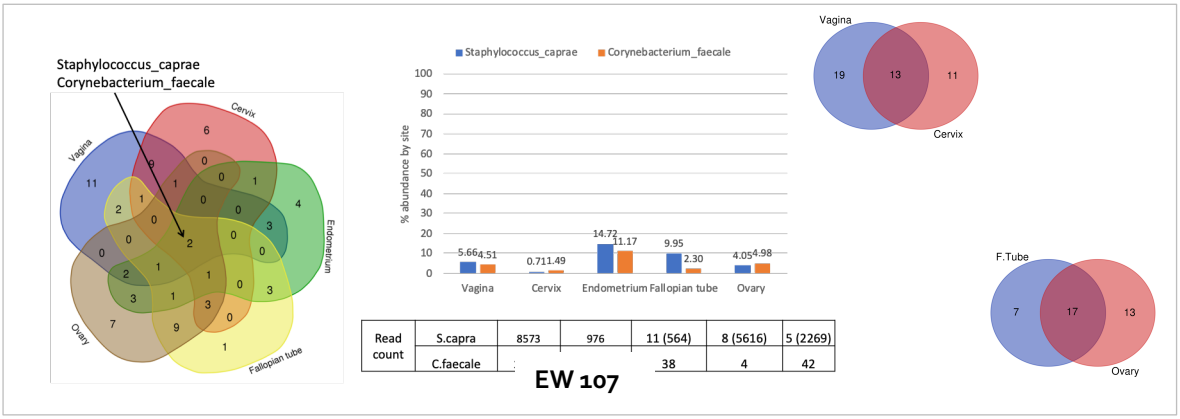
Read count	F.magna	8573	976	11 (286)	8 (373)	5 (424)
	P.unclassified			38	4	42

EW 99



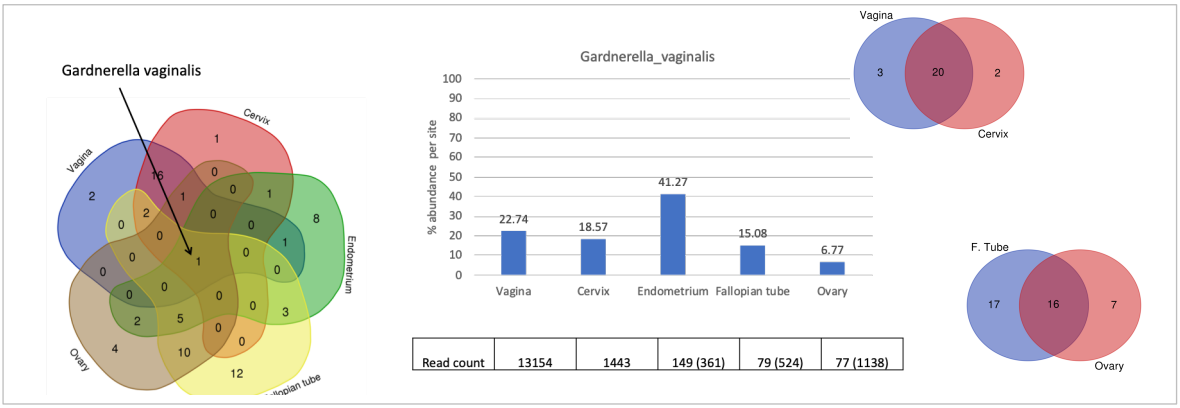
Read count	G.vaginalis	7 (83)	4 (87)	6 (59)	4 (157)	16 (97)
	L.crispatus	22	2	8	4	3
	L.gasseri	5				2

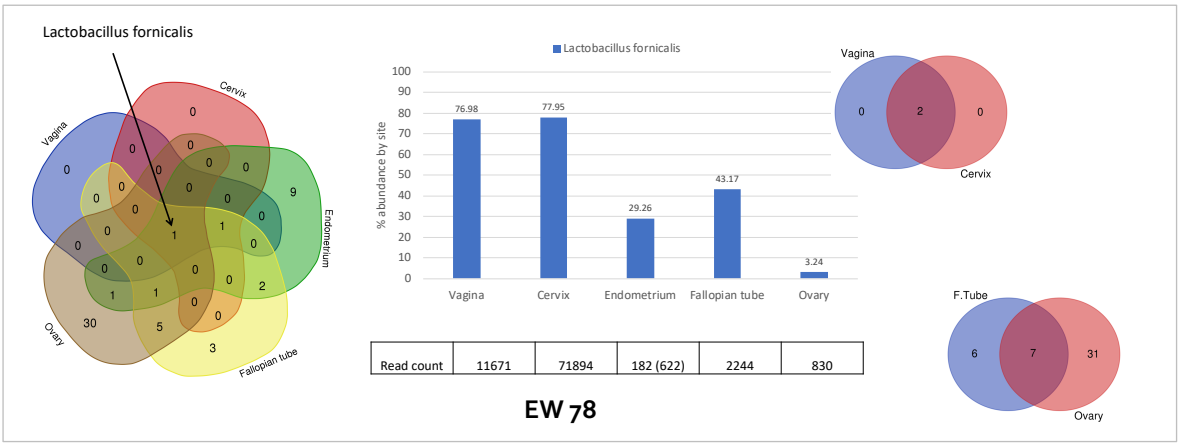
EW 105



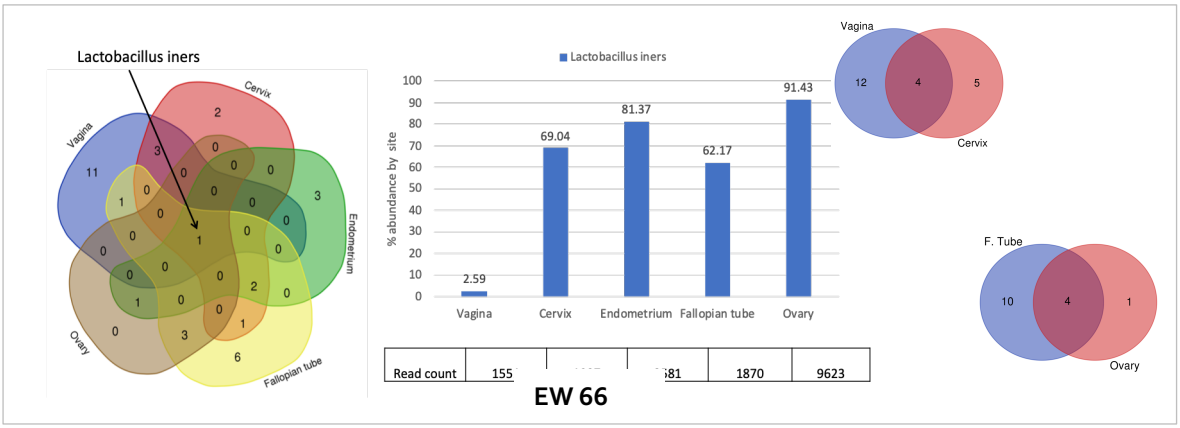
Read count	S.capra	8573	976	11 (564)	8 (5616)	5 (2269)
	C.faecale			38	4	42

EW 107

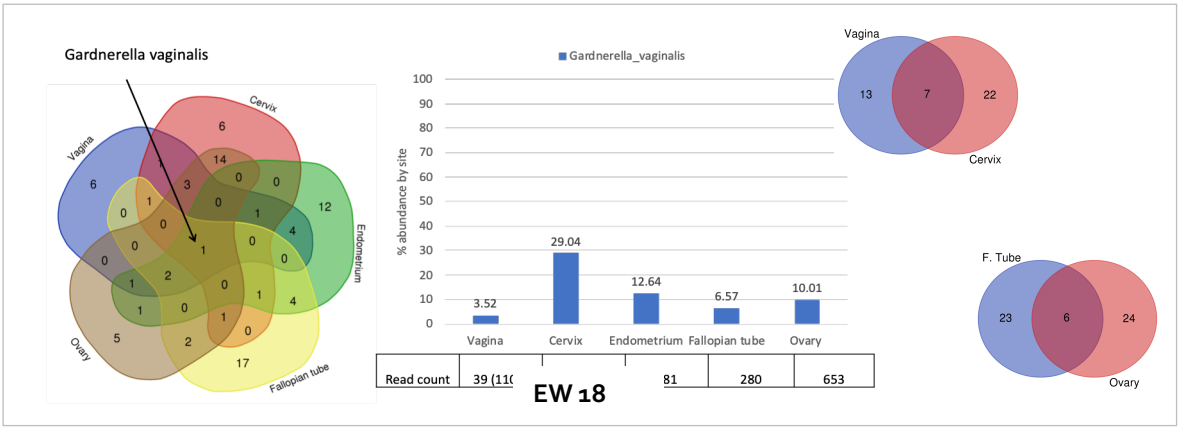




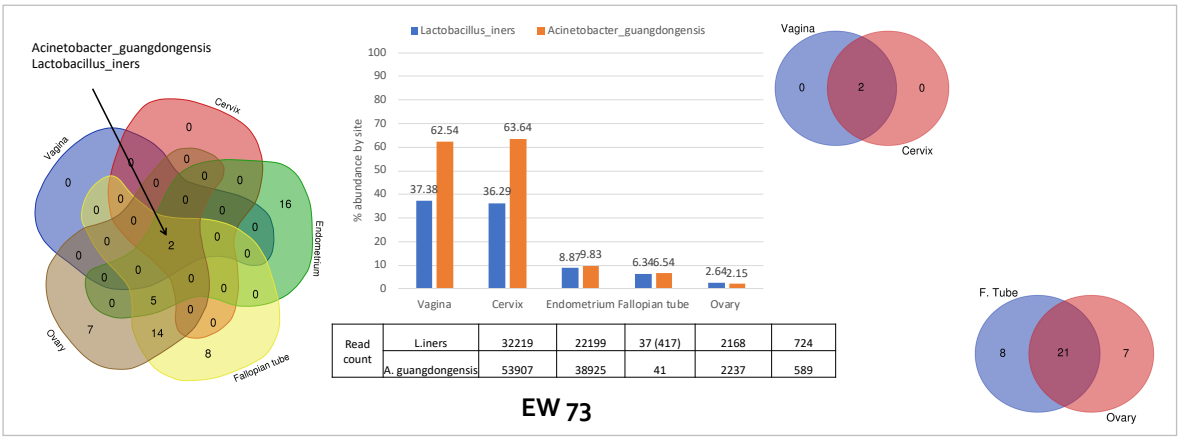
EW 78



EW 66

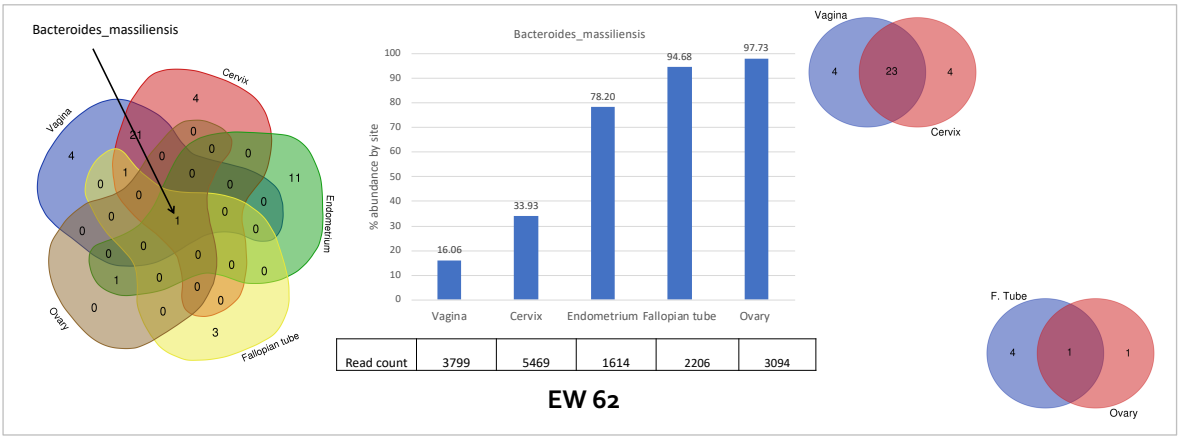


EW 18



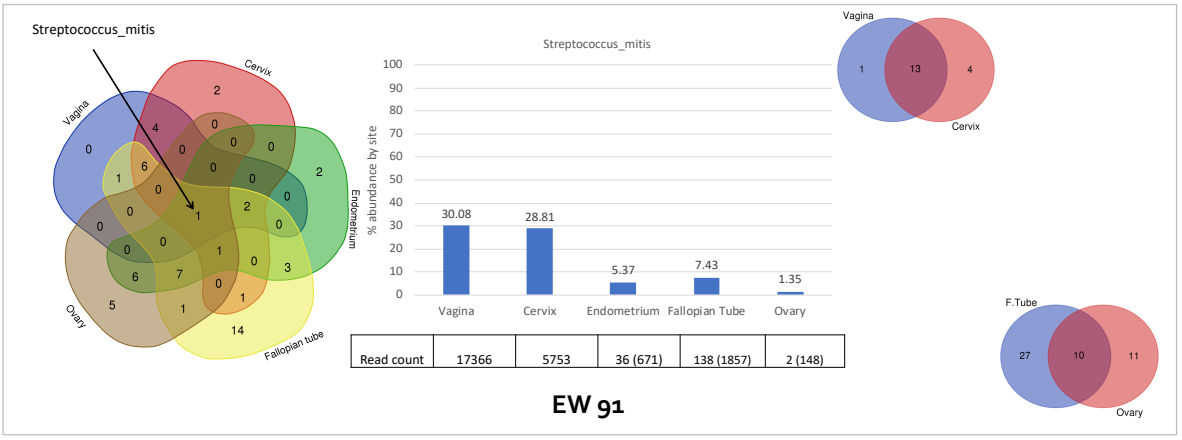
		Vagina	Cervix	Endometrium	Fallopian tube	Ovary
Read count	L.iners	32219	22199	37 (417)	2168	724
	A. guangdongensis	53907	38925	41	2237	589

EW 73



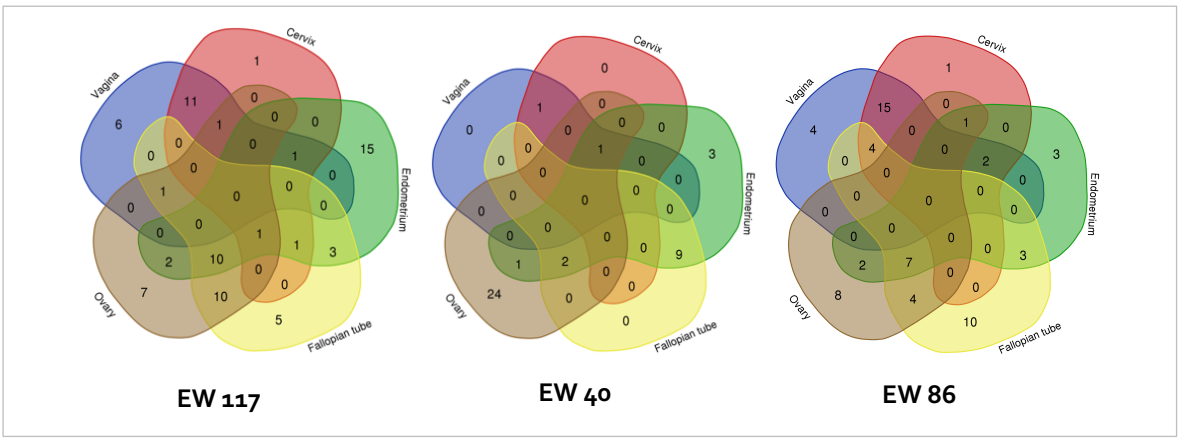
	Vagina	Cervix	Endometrium	Fallopian tube	Ovary
Read count	3799	5469	1614	2206	3094

EW 62



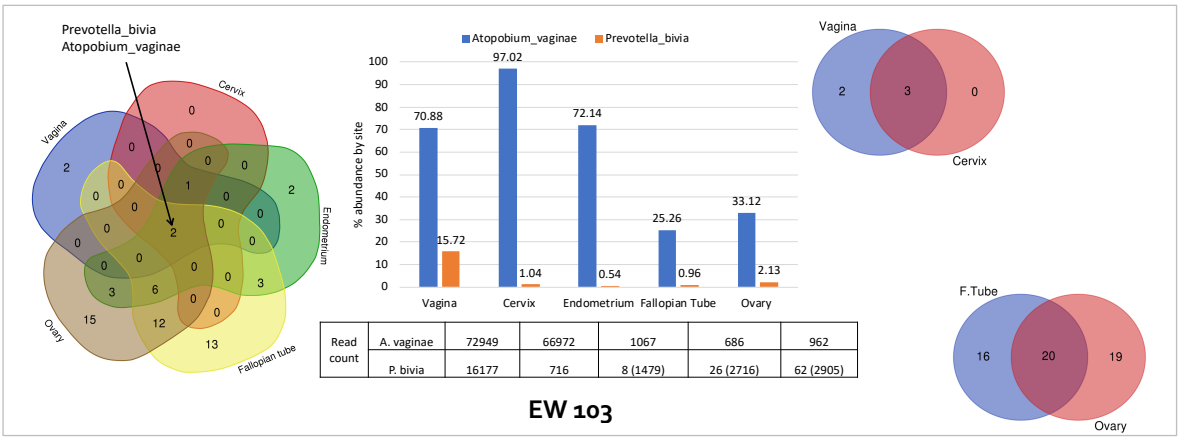
	Vagina	Cervix	Endometrium	Fallopian Tube	Ovary
Read count	17366	5753	36 (671)	138 (1857)	2 (148)

EW 91



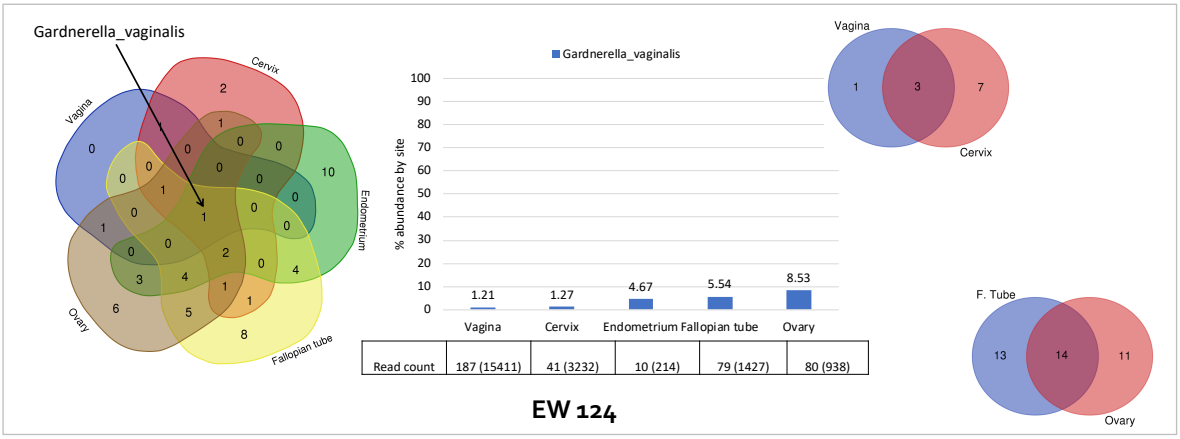
EW 117 **EW 40** **EW 86**

Benign samples



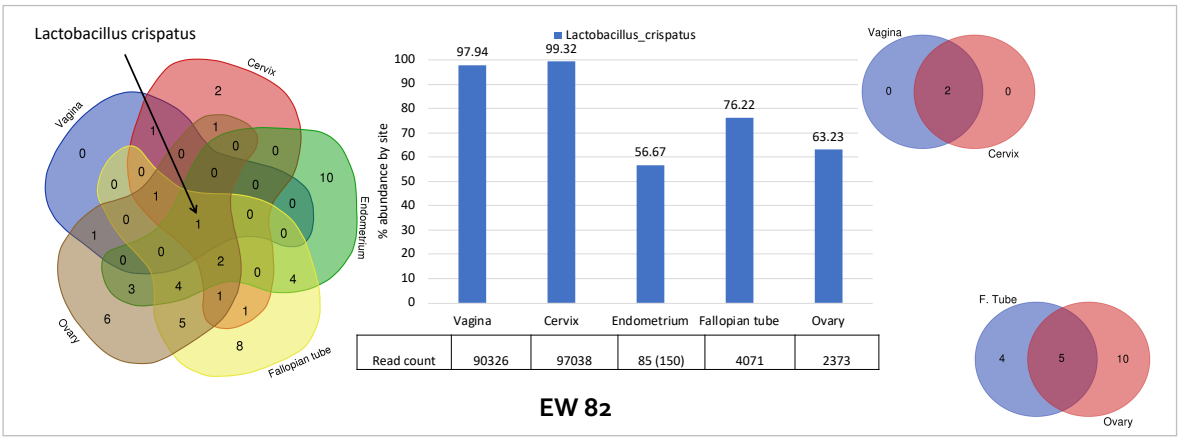
Read count	A. vaginae	72949	66972	1067	686	962
	P. bivia	16177	716	8 (1479)	26 (2716)	62 (2905)

EW 103



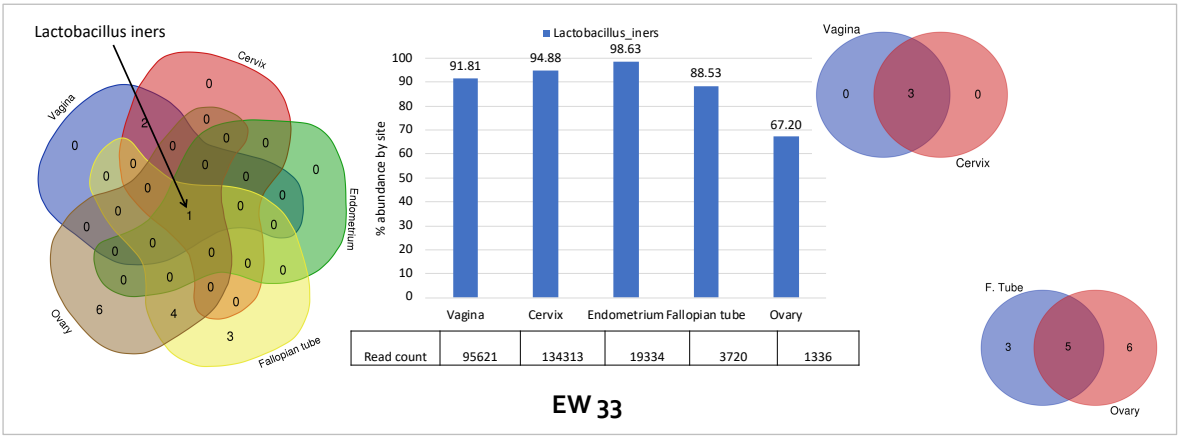
Read count	187 (15411)	41 (3232)	10 (214)	79 (1427)	80 (938)

EW 124



Read count	90326	97038	85 (150)	4071	2373

EW 82



Read count	95621	134313	19334	3720	1336

EW 33

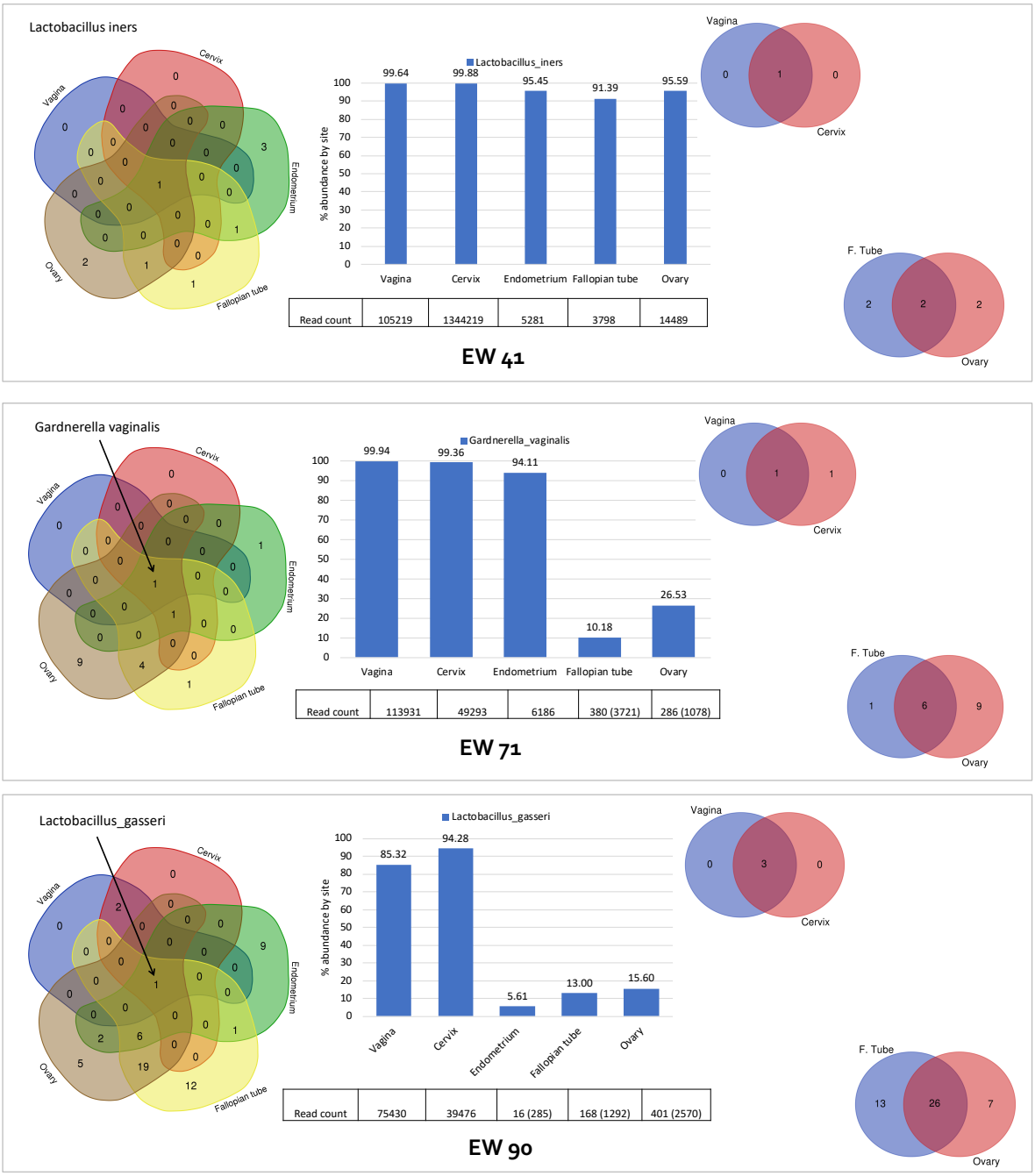
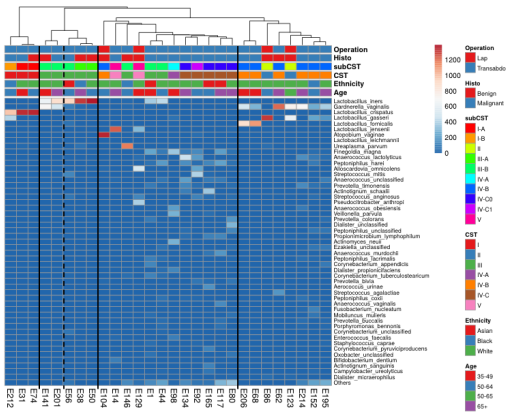


Figure S1 Common bacterial species detected at each anatomical site in the reproductive tract in (a) the malignant cohort, (b) the benign cohort. The large multi part Venn diagram illustrates how many bacterial species are shared between each anatomical space, with the 'common' microbe identified. The bar chart represents the percentage relative abundance of the common microbe in each space, with the actual read count in the table below the chart. The number of shared species between the LGT spaces, and the UGT spaces are also represented by the smaller Venn diagrams.

Chapter 4- Hierarchal clustering analysis by anatomical site at species level

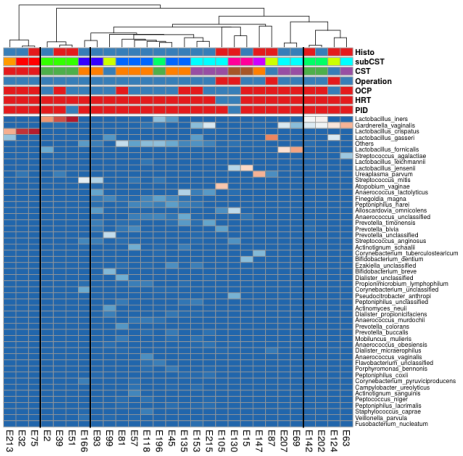
Cervix incl Val CST (Species)



Cervix	L.crispatus dominant (Group 1) n/N (%)	L. iners dominant (Group 2) n/N (%)	High diversity (Group 3) n/N (%)	G.vaginalis, L. fornicalis L.gasseri (Group 4) n/N (%)	Total n/N (%)
Benign	1/10 (10%)	3/10 (30%)	3/10 (30%)	3/10 (30%)	10/10 (100)
Cancer	2/18 (11.1)	2/18 (11.1)	9/18 (50)	5/18 (27.8)	18/18 (100)
Total	3/28 (10.7)	5/28 (10.7)	12/28 (42.9)	8/28 (28.6)	28/28 (100)
*Chi-squared (1.90)					p=.59279

Cervix	CST I-III, V (Lactobacillus dominant) N/n (%)	CST IV (Low lactobacillus) n/N (%)	Total n/N (%)
Benign	8/10 (80)	2/10 (20)	10/10 (100)
Cancer	7/18 (38.9)	11/18 (61.1)	18/18 (100)
Total	15/28 (53.6)	13/28 (46.4)	28/28 (100)
	*Fishers exact		p=.0546

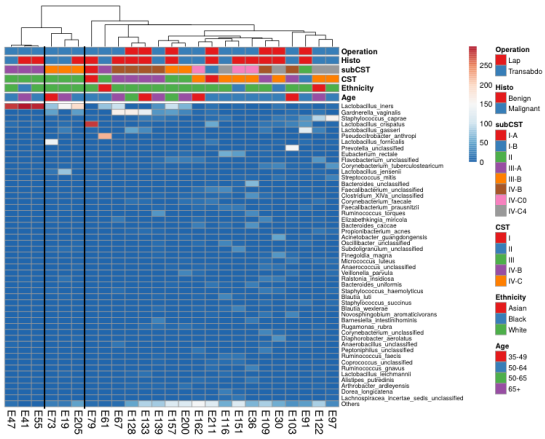
Vagina incl Valencia CSTs (Species)



Vagina	L.crispatus dominant (Group 1) n/N (%)	L. iners dominant (Group 2) n/N (%)	High diversity (Group 3) n/N (%)	L. iners, G.vaginalis, L. fornicalis (Group 4) n/N (%)	Total n/N (%)
Benign	1/10 (10%)	2/10 (20%)	4/10 (40%)	3/10 (30%)	10/10 (100)
Cancer	2/18 (11.1)	1/18 (5.6)	12/18 (66.7)	3/18 (16.7)	18/18 (100)
Total	3/28 (10.7)	3/28 (10.7)	16/28 (57.1)	6/28 (21.4)	28/28 (100)
*Chi-squared (2.59)					P=.45879

Vagina	CST I-III, V (Lactobacillus dominant) N/n (%)	CST IV (Low lactobacillus) n/N (%)	Total n/N (%)
Benign	7/10 (70)	3/10 (30)	10/10 (100)
Cancer	7/18 (38.9)	11/18 (61.1)	18/18 (100)
Total	14/28 (50)	14/28 (50)	28/28 (100)
	*Fishers exact		P=.2365

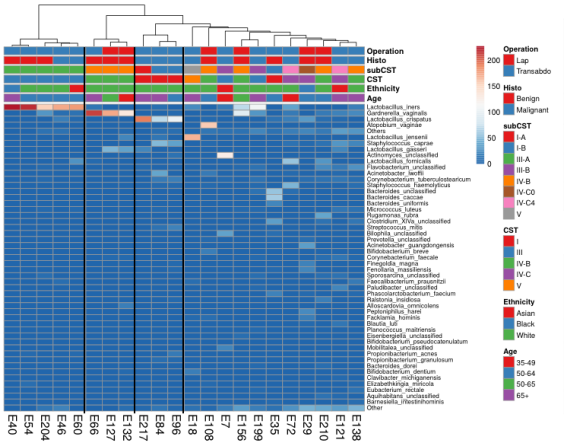
Endo Low incl Val CST (Species)



Endo Low	L. Iners dominant (Group 1) n/N (%)	L.iners/ G.vaginalis dominant (Group 2) n/N (%)	High diversity (Group 3) n/N (%)	Total n/N (%)
Benign	2/14 (14.3)	1/14 (7.1)	11/14 (78.6)	14/14 (100)
Cancer	1/11 (9.1)	2/11 (18.2)	8/11 (72.7)	11/11 (100)
Total	3/25(12)	3/25 (12)	19/25 (76)	25/25 (100)
*Chi-squared =0.79				p=.67309

Endo Low	CST I-III, V (Lactobacillus dominant) N/n (%)	CST IV (Low lactobacillus) n/N (%)	Total n/N (%)
Benign	7/14 (50)	7/14 (50)	14/14 (100)
Cancer	5/11 (45.5)	6/11 (54.5)	11/11 (100)
Total	12/25 (48)	13/25 (52)	25/25 (100)
	*Fishers exact		p=1

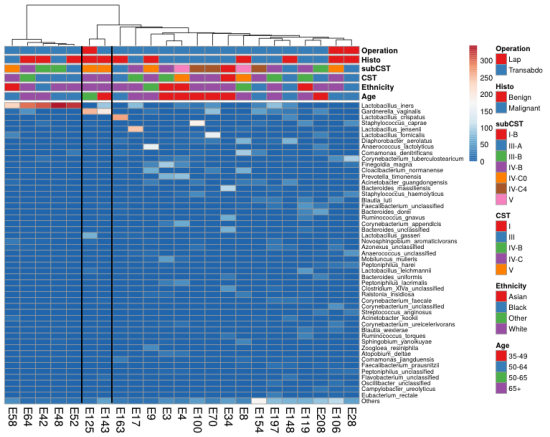
Endo High incl Val CST (Species)



Endo High	L.iners dominant (Group 1) n/N (%)	G.vaginalis dominant (Group 2) n/N (%)	L.Crispatus dominant (Group 3) n/N (%)	High diversity (Group 4) n/N (%)	Total n/N (%)
Benign	3/11 (27.3)	3/11 (27.3)	0/11 (0)	5/11 (45.5)	11/11 (100)
Cancer	2/11 (18.2)	0/11 (0)	3/11 (27.3)	6/11 (54.5)	11/11 (100)
Total	5/22 (22.7)	3/22 (13.6)	3/22 (13.6)	11/22 (50)	22/22 (100)
*Chi-squared = unable due to 0					p=.

Endo High	CST I-III, V (Lactobacillus dominant) N/n (%)	CST IV (Low lactobacillus) n/N (%)	Total n/N (%)
Benign	4/11 (36.4)	7/11 (63.6)	11/11 (100)
Cancer	8/11 (72.7)	3/11 (27.3)	11/11 (100)
Total	12/22 (54.5)	10/22 (45.5)	22/22 (100)
	*Fishers exact		p=1984

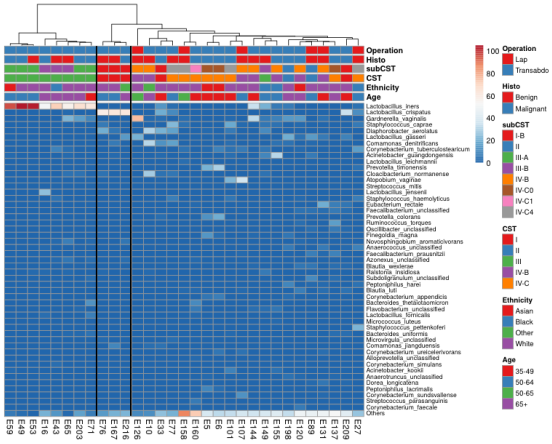
Fall Tube incl Val CST (Species)



Fall Tube	L.iners dominant (Group 1) n/N (%)	G.vaginalis dominant (Group 2) n/N (%)	High diversity (Group 3) n/N (%)	Total n/N (%)
Benign	3/11 (27.3)	2/11 (18.2)	6/11 (54.5)	11/11 (100)
Cancer	2/12 (16.7)	0/12 (0)	10/12 (83.3)	12/12 (100)
Total	5/23 (21.7)	2/23 (8.7)	16/23 (69.6)	23/23 (100)
*Chi-squared = unable due to 0				p=.

Fall tube	CST I-III, V (Lactobacillus dominant) N/n (%)	CST IV (Low lactobacillus) n/N (%)	Total n/N (%)
Benign	6/11 (54.5)	5/11 (45.5)	11/11 (100)
Cancer	5/12 (41.7)	7/12 (58.3)	12/12 (100)
Total	11/23 (47.8)	12/23 (52.2)	23/23 (100)
	*Fishers exact		p=.6843

Ovary incl Val CST (Species)



Ovary	L.iners dominant (Group 1) n/N (%)	L. crispatus dominant (Group 2) n/N (%)	High diversity (Group 3) n/N (%)	Total n/N (%)
Benign	3/15 (20)	2/15 (13.3)	10/15 (66.7)	15/15 (100)
Cancer	5/16 (31.3)	1/16 (6.2)	10/16 (62.5)	16/16 (100)
Total	8/31 (25.8)	3/31 (9.7)	20/31 (64.5)	31/31 (100)
*Chi-squared =				p=.66968

Ovary	CST I-III, V (Lactobacillus dominant) N/n (%)	CST IV (Low lactobacillus) n/N (%)	Total n/N (%)
Benign	7/15 (46.7)	8/15 (53.3)	15/15 (100)
Cancer	9/16 (56.3)	7/16 (44.7)	16/16 (100)
Total	16/31 (51.6)	15/31 (48.4)	31/31 (100)
	*Fishers exact		p=.7244

Figure S2 Hierarchal clustering analysis of samples by anatomical site at species level.

Chapter 5- α and β- diversity results by anatomical site at both species and genus level

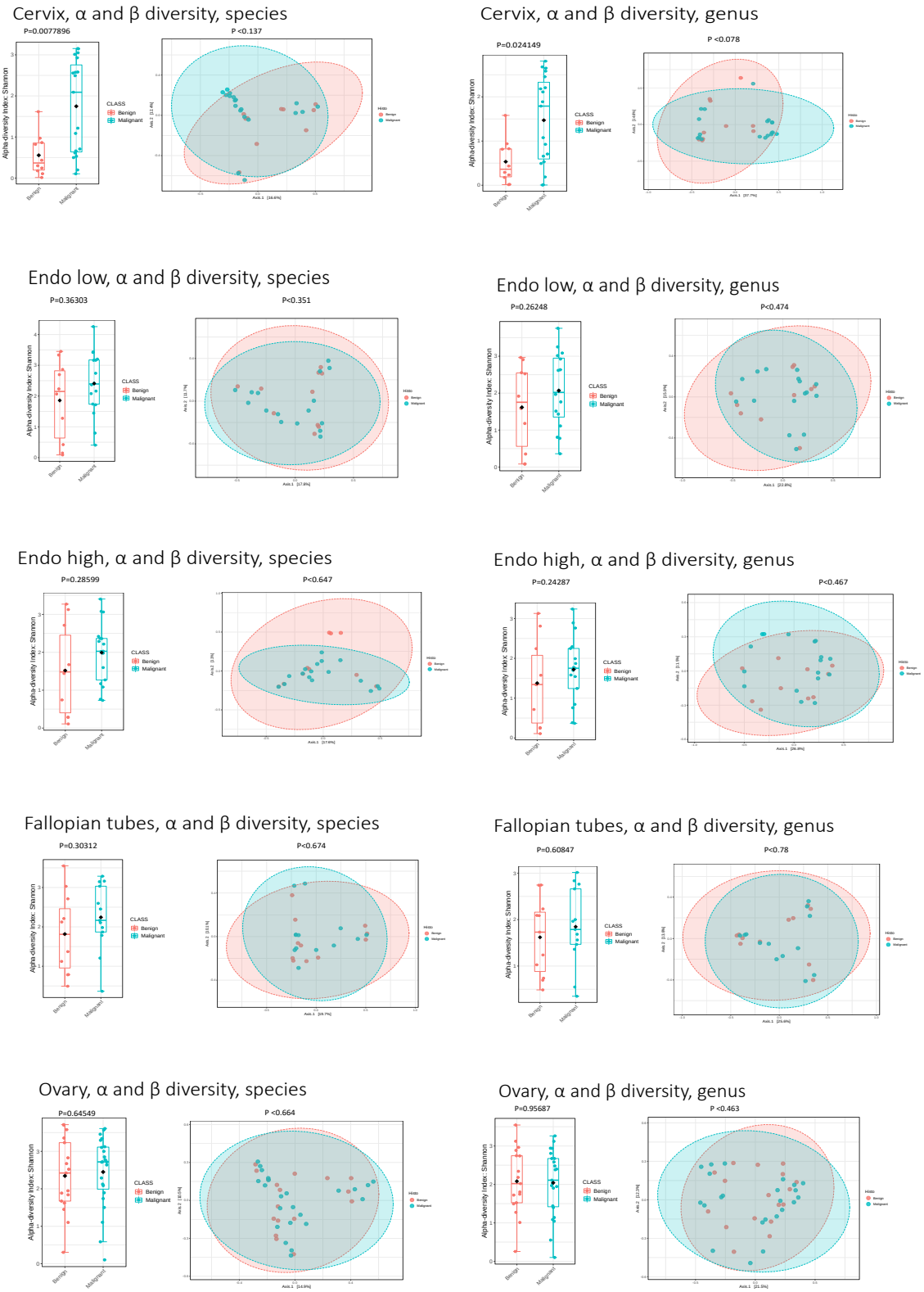


Figure S3 α and β diversity of anatomical sites at species and genus level.

PATIENT CONSENT FORM

The consent form used to recruit participants into this study is shown below.

Imperial College Healthcare NHS
NHS Trust

AFFIX PATIENT STICKER HERE

Hammersmith Hospital
Du Cane Road
London
W12 0HS
Tel: 020 8383 1000
www.imperial.nhs.uk

Patient Consent Form

Version 8.0 (01/01/2017)

Metabonomics and Integrative Biology for Gynaecological Cancer

1	I confirm that I have read and understood the patient information sheet for he above study and have had the opportunity to ask questions.	☐
2	I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without my medical care or legal rights being affected.	☐
3	I understand that sections of any of my medical notes and the completed questionnaire may be looked at by responsible individuals or by regulatory authorities where it is relevant to my taking part in research. I give permission for these individuals to have access to my records.	☐
4	I agree to allow a sample of tissue obtained from my surgery / biopsy to be used in this research project. I understand that giving this sample is voluntary and that I am free to withdraw my approval for its use at any time without giving a reason and without my medical care or legal rights being affected.	☐
5	I agree to give blood, swabs, stool, and urine samples for research in this project. I understand that giving these samples is voluntary and that I am free to withdraw my approval for use of the sample at any time without giving a reason and without my medical care or legal rights being affected.	☐

6 I agree that the samples may be used to identify the presence of bacteria. ☐

7 I agree that the samples I have given and the information gathered about me can be stored at Imperial College for possible use in future studies to further the understanding of gynaecological cancers. ☐

8 I understand that some of these projects may be carried out by researchers other than Imperial College. I understand that some of these projects may involve international collaborations. ☐

9 I understand that I shall not benefit financially if future research leads to the development of new treatments or medical tests. ☐

10 I understand that confidentiality will be maintained as all information about me will be anonymous. ☐

11 I agree to take part in the above study. ☐

Name of Patient	Date	Signature
Name of Person taking consent	Date	Signature

PATIENT INFORMATION LEAFLET

Patient information leaflet provided to all participants recruited from either the gynaecology clinic or the gynaecology ward at Queen Charlotte's and Chelsea Hospital, Imperial College Healthcare NHS Trust.

Imperial College London

Imperial College Healthcare NHS Trust

Hammersmith Hospital
Du Cane Road
London
W12 0HS
Tel: 020 8383 1000
www.imperial.nhs.uk

Patient Information Sheet

You are being invited to take part in a research study. Before you decide it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with friends, relatives, and members of hospital staff. Ask us if there is anything that is not clear or if you would like more information. Please take time to decide whether or not you wish to take part. If you do decide to take part, please let us know beforehand if you have been involved in any other study during the last year. If you decide not to take part your future care will not be affected by your decision. You are free to withdraw at any time without explanation and your future care will not be affected. Thank you for reading this.

What is the purpose of the study?

We at Imperial College London are constantly trying to improve our knowledge about gynaecological cancer and other gynaecological conditions. This project is designed to see the differences that can be detected in people with different gynaecological conditions in order to achieve a comprehensive understanding of these underlying causes of the diseases. Detailed and more technical information about the projects is available from clinic staff. Please note that you are not taking part in a clinical treatment or drug trial.

Why have I been chosen?

You have been chosen as you may be having investigations for a gynaecological condition.

Do I have to take part?

No. Taking part in this study is voluntary. Whatever your decision, it will not affect your current or future medical care or rights in any way. If you do decide to take part you may withdraw from the study at any time without having to give a reason, and it will make no difference to your medical care or treatment.

If you decide to withdraw from the study, tissue samples or data may already be retained and used for the purposes for which consent has already been given, provided they are effectively anonymised and no longer identifiable to the research team or any other persons to whom access will be given. I understand that if I lose capacity to consent whilst I am a participant in the study, all previous data and tissue collected with consent will be retained and used, but no further data collection or procedures will be carried out.

What will happen to me if I take part?

When you come into the hospital for your appointment or procedure, we will ask you if you would agree to complete a questionnaire with details from your medical, personal, and family history and donate blood, urine, and tissue (or swabs) from your genital tract (including your vagina, cervix, uterus, fallopian tubes, and ovaries) for research purposes. We would like to study these samples. It will take you approximately 20 minutes to fill in the questionnaire. We will collect an extra 20-40 mls (4-8 teaspoons) of blood along with a urine sample. We will collect these on the day of your appointment/procedure, while you are an inpatient (if you are having surgery). Tissue and swabs from your genital tract will be taken on the day of your appointment or procedure. The samples will be taken when you are asleep under a general anesthetic. The clinical information we have and will continue to collect in your hospital records may be analysed.

We at Imperial College London, often collaborate with international institutions in order to gain as much information as possible about gynaecological conditions. If you consent to this study it is possible that some of your samples may be sent to other international centres for further research. All information and samples will be coded in such a way as to make it impossible for you to be identified personally.

You will not be paid for participating in this study. We may be able to reimburse you for reasonable expenses (parking and travel) related to your study visits, over and above visits required if you were not participating in this study.

What are the possible benefits of taking part?

There are no direct benefits to you if you take part in this study. The projects underway now and in the future may shed light on what is going on in gynaecological conditions and cancers. If you do decide to participate we hope that the knowledge gained from your participation may benefit patients with gynaecological conditions in the future.

What are the possible disadvantages and risks of taking part?

There are no major risks from taking part in this study. The blood tests can cause some discomfort, pain, or a bruise.

Will my taking part in this study be kept confidential?

All of your personal details will be kept confidential. Your samples will be held at Imperial College and will only be identified by a unique number. If your samples are sent to a centre outside of Imperial College, they will also be identified only by a unique code.

Who is organising the research?

This study is organised by the staff of Imperial College NHS Trust and Imperial College London. This research may be part of a PhD or MD study proposal of one of the investigators.

Who has reviewed the study?

The West of Scotland REC 5 Research Ethics Committee has reviewed this study as well as it being independently peer reviewed

What if something goes wrong?

Imperial College London holds insurance policies, which apply to this study. If you experience serious and enduring harm or injury as a result of taking part in this study, you may be eligible to claim compensation without having to prove that Imperial College is at fault. This does not affect your legal rights to seek compensation. If you are harmed due to someone's negligence, then you may have grounds for a legal action. Regardless of this, if you wish to complain, or have any concerns about any aspect of the way you have been treated during the course of this study then you should immediately inform the Investigator **Dr Maria Kyrgiou** on 020 8383 3268. The normal NHS complaints mechanisms are also available to you. If you are still not satisfied with the response, you may contact the Imperial PALS (Patient Advice Liaison Service) on 0208 3130088.

Contact for Further Information

If you want more information, before or after you return your form, you can phone Dr Maria Kyrgiou on 020 8383 3268.

Imperial College London

Imperial College Healthcare NHS Trust

Participant's Identifiable information: Today's date: ____/____/_____

Height (cm): _____

Weight (kg): _____

BMI: _____

Section A: Demographic Data

A1. Which of these best describes your ethnic group? *(Please mark only one.)*

White

□ 1. British

□ 2. Irish

□ 3. Gypsy or Irish Traveller

□ 4. Any other White background, Please specify

_____

Mixed/multiple ethnic groups

□ 5. White and Black Caribbean

□ 6. White and Black African

□ 7. White and Asian

□ 8. Any other Mixed/Multiple ethnic background, please specify

_____

Asian/Asian British

□ 9. Indian

□ 10. Pakistani

□ 11. Bangladeshi

□ 12. Chinese

□ 13. Any other Asian background, please specify

_____

Black/African/Caribbean/Black British

□ 14. African

□ 15. Caribbean

□ 16. Any other background, please specify_____

Other ethnic groups

□ 17. Arab

□ 18. Any other ethnic group, please specify

A2. What is the highest level of education that you have completed?

□ No formal education □ GCSE's or equivalent □ A-levels or equivalent

□ Degree/Higher degree □ Other, please specify _____

A3. What is your marital status?

□ Single □ Married/civil partnership/living together □ Divorced/separated

□ Widowed

Section B: Reproductive history

This section asks about pregnancies. Please answer, even if you have never been pregnant.

B1. Have you ever been pregnant?

□ No

□ Yes – Total number of pregnancies _____ (Include live births, stillbirths, terminations, miscarriages, ectopic pregnancies)

Number of live births _____ (Include live births & stillbirths only)

B2. Have you ever tried to become pregnant for one full year, without becoming pregnant?

□ No □ Yes □ Not sure

B3. If you had difficulty getting pregnant, which of these best describes the reason?

□ Problem with ovaries or hormones □ Problem with uterus or cervix □ Male partner fertility problem

□ Problems with fallopian tubes □ No problem found □ Not sure

B4. Did your doctor ever prescribe medication to help you become pregnant?

□ No □ Yes □ Not sure

Section C: Menstrual history

This section asks questions about your menstrual cycle, menstrual history and menopause.

C1. Have you ever had a menstrual period?

□ No

□ Yes ® At what age did you have your first menstrual period? _____ years

C2. Which best describes the pattern of your menstrual cycle over most of your life?

□ Regular (could usually predict when I would occur) □ Irregular (not predictable)

C3. Which best describes your current menstrual status?

□ I still have regular periods (Go to **C5**)

□ I still have irregular periods (Go to **C5**)

□ Yes: No menstrual periods ® at what age was your last period? _____ years

C4. If your menstrual periods ceased PERMANENTLY, which best describes why your periods stopped?

□ Natural menopause □ Surgery (uterus+/-ovaries removed) □ Radiation

□ Medication/drug therapy □ Other, please specify _____

C5. Have you ever had a hysterectomy (surgery to remove uterus/womb)?

□ No

□ Yes –At what age did you have your hysterectomy? _____ years

C6. Have you ever had surgery to remove one/both ovaries?

□ No

□ Yes, one ovary removed ® at what age did you have one ovary removed? _____ years

□ Yes, both ovaries removed ® at what age did you have both ovaries removed? _____ years

C7. Have you ever had surgery to tie your fallopian tubes (female sterilization)?

□ No

□ Yes ® At what age did you have your tubes tied? _____ years

Section D: Oral Contraceptive Use

This section asks questions about current and past use of oral contraceptives ('The Pill').

D1. Have you ever taken oral contraceptives ('The Pill')?

(This includes use for reasons other than contraception)

□ No

□ Yes –At what age did you start taking oral contraceptives? _____ years

– Total duration of oral contraceptive use (minus any breaks) _____ years

– Are you currently taking oral contraceptives □ Yes □ No ® age when stopped use ___ years

Section E: Hormone Replacement Therapy Use

This section asks questions about current and past use of hormone replacement therapy ('HRT').

E1. Have you ever taken hormone replacement therapy, or medication containing female hormones for relief of menopausal symptoms, irregular periods or to prevent osteoporosis (bone thinning)?

□ No (Go to section F)

□ Yes

Hormone replacement therapy may contain oestrogen alone, progesterone alone, or a combination of the two. Please provide information, based on the type/types of HRT that you are taking/have taken.

E2. Have you ever taken hormone replacement therapy containing oestrogen/oestradiol *alone*?

(e.g. Climaval®, Elleste-Solo®, Evorel®, Premarin®, Progynova®)

□ No

□ Yes –At what age did you start taking these tablets? _____ years

- Total duration of this type of HRT use _____ years
- Are you currently taking this type of HRT□ Yes □ No ® age when stopped use ___ years

E3. Have you ever taken hormone replacement therapy containing progesterone/progestogen *alone*?

(e.g. Climanor®, Utrogestan®, Provera, Norethisetrone)

□ No

□ Yes – At what age did you start taking these tablets? _____ years

- Total duration of this type of HRT use _____ years
- Are you currently taking this type of HRT□ Yes □ No ® age when stopped use ___ years
- How many days per month do/did you take these tablets? _____ days

E4. Have you ever taken hormone replacement therapy containing oestrogen *and* progesterone?

(eg. Climagest®, Elleste-Duet®, Femoston®, Kliofem®, Kliovance®, Premique®, Prempak-C®)

□ No

□ Yes –At what age did you start taking these tablets? _____ years

- Total duration of this type of HRT use _____ months/ _____ years
- Are you currently taking this type of HRT□ Yes □ No ® age when stopped use ___ years
- What is the name of the brand you took for the longest? _____

E5. Have you ever worn patches containing hormone replacement therapy?

(e.g. Elleste conti®, Estraderm®, Evorel®, FemSeven®)

□ No

□ Yes –At what age did you start wearing these patches? _____ years

- Total duration of this type of HRT use _____ months/ _____ years
- Are you currently using HRT patches □ Yes □ No ® age when stopped use ___ years

E6. Have you ever used any other form of hormone replacement therapy? (Mark all that apply)

□ Vaginal cream/gel □ Vaginal ring □ Progesterone IUD

□ Other, please specify _____

Section F: Body Mass

This section asks questions about height and weight. If you are not sure, please estimate.

F1. What is your height? _____ feet, _____ inches or _____ centimetres

F2. What was your weight at the following times? (If you were pregnant at the time, please state weight immediately prior to pregnancy)

12 months ago? _____ stone or _____kg or _____lbs

10 years ago? _____ stone or _____kg or _____lbs

at 18 years of age? _____ stone or _____kg or _____lbs

F3. On which area of your body do you tend to gain weight?

□ Chest/shoulders □ Waist/stomach □ Hips/thighs

□ Equally all over □ Don't gain weight □ Other, please specify

Section G: Social History

This section asks questions about cigarette smoking.

G1. Have you ever smoked more than 100 cigarettes or more, in your lifetime?

□ No

□ Yes – Have you ever smoked cigarettes regularly? (i.e., at least one per day, for 6 months or more)

□ No (Go to section H)

□ Yes – At what age did you start smoking regularly? _____ years

– Do you smoke currently? □ No ® at what age did you stop? _____ years

□ Yes

G2. How many cigarettes/day do/did you smoke on average? _____ cigarettes/day

G3. Do you drink coffee regularly (daily basis)? □ Yes □ No

On average, how many cups of ***caffeinated*** coffee do you drink each day? □ 1 □ 2 □ 3 □ >4

On average, how many cups of ***decaffeinated*** coffee do you drink each day? □ 1 □ 2 □ 3 □ >4

Section H: Medical History

Please tick 'yes' if a doctor has ever diagnosed you with any of these medical problems, and state your approximate age at diagnosis.

Condition	
H1a. Hyperthyroidism (overactive thyroid)?	□ No □ Yes ® Age at diagnosis _____ years
H1b. Hypothyroidism (underactive thyroid)?	□ No □ Yes ® Age at diagnosis _____ years
H1c. Adult onset diabetes (Type 2 diabetes mellitus)?	□ No □ Yes ® Age at diagnosis _____ years
H1d. Diabetes (Type 1 diabetes mellitus)?	□ No □ Yes ® Age at diagnosis _____ years
H1e. Hypertension (high blood pressure)?	□ No □ Yes ® Age at diagnosis _____ years
H1f. Endometriosis?	□ No □ Yes ® Age at diagnosis _____ years
H1g. Fibroid (non-cancerous tumors) of the uterus?	□ No □ Yes ® Age at diagnosis _____ years
H1h. Polycystic ovarian disease or Stein-Leventhal Syndrome?	□ No □ Yes ® Age at diagnosis _____ years
H1i. Other non-cancerous cysts or tumors of the ovary?	□ No □ Yes ® Age at diagnosis _____ years
H1j. Pelvic inflammatory disease (PID)?	□ No □ Yes ® Age at diagnosis _____ years
H1k. Breast cancer?	□ No □ Yes ® Age at diagnosis _____ years
H1l. Borderline ovarian cancer (ovarian tumor of low malignant potential or LMP)?	□ No □ Yes ® Age at diagnosis _____ years
H1m. Invasive ovarian cancer?	□ No

	□ Yes ® Age at diagnosis _____ years
H1n. Endometrial cancer (cancer of the lining of the uterus/womb)?	□ No □ Yes ® Age at diagnosis _____ years
H1o. Colon (bowel) cancer?	□ No □ Yes ® Age at diagnosis _____ years
H1p. Other types of cancer not listed above If yes, please specify _____	□ No □ Yes ® Age at diagnosis _____ years

Section I: Family History of Cancer

This section asks questions about the medical history of your family. Please only include *blood relatives*, with whom you share both biological parents and not adopted or step relatives.

I1. How many full sisters and/or full brothers do/did you have? ____ Sisters ____ Brothers

I2. Have many daughters and/or sons have you given birth to? _____ Daughters _____ Sons

I3. Has your mother, any full sisters or daughters ever been diagnosed with any type of cancer?

□ No

□ Yes ® □ Mother □ Sister □ Daughter □ All

Please specify the type of cancer

_____

How old she was when the cancer was first diagnosed

_____

I4. Has your father, any full brother or sons ever been diagnosed with any type of cancer?

□ No

□ Yes ® □ Father ☐ Brother ☐ Son ☐ All

Please specify the type of cancer

_____

How old he was when the cancer was first diagnosed _____

Section J: Medication History

This sections asks questions about over-the-counter medications you take/haven taken both as a once off and regularly. (This means at least once a week for one year, or more than 50 tablets in a one-year period).

J1. Have you ever taken metformin, insulin, or other diabetic medication, regularly?

□ No

□ Yes - What is the name of the mediation(s) _____

– At what age did you start taking this medication regularly? _____ years

– Total duration of regular use _____ years

– Are you currently taking this medication □ Yes □ No

If more than one diabetic medication, please specify details for each_____

_____

J2. Have you ever taken aspirin, or aspirin-containing medications, regularly?

□ No

□ Yes – At what age did you start taking this medication regularly? _____ years

– Total duration of regular use _____ years

– Are you currently taking this medication □ Yes □ No

J3. Have you ever taken paracetamol, or paracetamol-containing medications, regularly?

□ No

□ Yes – At what age did you start taking this medication regularly? _____ years

– Total duration of regular use _____ years

– Are you currently taking this medication □ Yes □ No

J4. Have you ever taken painkillers or anti-inflammatory drugs *other than* aspirin, or paracetamol-containing medications, regularly? (e.g. Feminax®, Neurofen®(ibuprofen), Naproxen, Indomethacin, Piroxicam)

□ No

□ Yes – At what age did you start taking this medication regularly? _____ years

– Total duration of regular use _____ years

– Are you currently taking this medication □ Yes □ No

J5. Have you ever taken COX-2 inhibitors (usually prescribed for arthritis) regularly?

(e.g. Celebrex® (celecoxib), Vioxx®(rofecoxib)

□ No

□ Yes – At what age did you start taking this medication regularly? _____ years

– Total duration of regular use _____ years

– Are you currently taking this medication □ Yes □ No

J6. Have you ever taken Tamoxifen regularly?

□ No

□ Yes – At what age did you start taking this medication regularly? _____ years

– Total duration of regular use _____ years

– Are you currently taking this medication □ Yes □ No

J7. Have you ever taken Raloxifene (Evista®) regularly?

□ No

□ Yes – At what age did you start taking this medication regularly? _____ years

– Total duration of regular use _____ years

– Are you currently taking this medication □ Yes □ No

J8. Have you ever taken bisphosphonates (drugs to prevent osteoporosis/bone thinning) regularly?

(e.g. Fosamax® (alendronate), Actonel®(Risedronate),

□ No

□ Yes – At what age did you start taking this medication regularly? _____ years

– Total duration of regular use _____ years

– Are you currently taking this medication □ Yes □ No

J9. Have you taken antibiotics, antifungals e.g. Canestan™/ clotrimazole, or probiotics e.g. Yakult™, BioKult™ in the past 4 weeks? Please document all antibiotics etc. if there are more than one.

□ No

_____ □ Yes - What is the name of the antibiotic/antifungal/probiotic (s)?

- _____ For how long did you take it? days
- What dose did you take? _____mg
- What route was the medication given (circle as appropriate)?

  Oral Vaginal Cream/ointment Via drip Intra- Muscular
- What was the reason for the medication? _____

J10. Have you used any vaginal pessaries/ tablets/ creams or ointments in the past 4 weeks?

□ No

□ Yes - What did you use? _____

- For long? _____

J11. Do you practice vaginal douching (internal washing of the vagina)?

□ No

□ Yes - How often: Once per day □ more frequently: _____

Once per week □ twice per week □ Any other _____

J12. Have you had vaginal intercourse in the past 2 weeks?

□ No

□ Yes

THANK YOU FOR YOUR TIME!

PUBLICATIONS

Published:

- **Whelan E,** Kalliala I, Semertzidou A, Raglan O, Bowden S, Kechagias K, Markozannes G, Cividini S, McNeish I, Marchesi J, MacIntyre D, Bennett P, Tsilidis K, Kyrgiou M. Risk Factors for Ovarian Cancer: An Umbrella Review of the Literature. Cancers (Basel). 2022 May 30;14(11):2708. doi: 10.3390/cancers14112708. PMID: 35681688; PMCID: PMC9179274.

Other publications arising from work related to this thesis:

- Paraskevaidi M, Cameron S.J.S, **Whelan E,** Bowden S, Tzafetas M, Mitra A, Semertzidou A, Athanasiou A, Bennett P.R, macIntryre D.A, Takats Z, Kyrgiou M. Laser-assisted rapid evaporative ionization mass spectrometry (LA-REIMS) as a metabolomics platform in cervical cancer screening. eBioMedicine. 2020 September 25; 60:103017. doi: https://doi.org/10.1016/j.ebiom.2020.103017. PMID: 32980699; PMCID: PMC7522750

PRESENTATIONS

Oral

Regional

- Risk factors for ovarian cancer and the microbiome. Work in Progress Series, Imperial College London, London UK 2019.

Poster

International

- **E Whelan** 'Risk factors for ovarian cancer; an umbrella review of the literature' Royal College of Obstetrics and Gynaecology World Congress, 2021

- **E Whelan** 'Risk factors for ovarian cancer; an umbrella review of the literature' European Society of Gynaecological Oncology, Austria 2019

National

◊ E **Whelan** 'Risk factors for ovarian cancer; an umbrella review of the literature' British Gynaecological Cancer Society Annual Scientific Meeting, London 2021

◊ **Whelan E** 'Automated High-Throughput Cytology using Rapid Evaporative Ionization Mass Spectrometry (REIMS)' British Society of Colposcopy and Cervical Pathology conference abstract, Manchester 2018