A Prospective Analysis of Circulating Plasma Metabolites Associated with Ovarian Cancer Risk

Oana A. Zeleznik1, A. Heather Eliassen1,2, Peter Kraft2, Elizabeth M. Poole1, Bernard A. Rosner1, Sarah Jeanfavre3, Amy A. Deik3, Kevin Bullock3, Daniel S. Hitchcock3, Julian Avila-Pacheco3, Clary B. Clish3, and Shelley S. Tworoger2,4

ABSTRACT

Ovarian cancer has few known risk factors, hampering identification of high-risk women. We assessed the association of prescreening plasma metabolites (N = 420) with risk of epithelial ovarian cancer, including both borderline and invasive tumors. A total of 252 cases and 252 matched controls from the Nurses’ Health Studies were included. Multivariable logistic regression was used to test ORs and 95% CI for the association of plasma metastable metabolite levels, using permutation-based Westfall and Young approach to account for testing multiple correlated hypotheses. Weighted gene coexpression network analysis (WGCNA; n = 10 metabolite modules and metabolic set enrichment analysis (n = 23 metabolite classes) were also evaluated. An increase in pseudouridine levels from the 10th to the 90th percentile associated with a 2.3-fold increased risk of ovarian cancer (OR = 2.36; 95% CI, 1.48–4.45; P = 0.001/adjusted P = 0.15); a similar risk estimate was observed for serous/poorly differentiated tumors (n = 176 cases; comparable OR = 2.38; 95% CI, 1.33–4.32; P = 0.004/adjusted P = 0.53). For nonserous tumors (n = 34 cases), pseudouridine and C36:2 phosphatidylcholine plasmalogen had the strongest statistical associations (OR = 9.84; 95% CI, 2.89–37.82; P < 0.001/adjusted P = 0.07; and OR = 0.11, 95% CI, 0.03–0.35; P < 0.001/adjusted P = 0.06, respectively). Five WGCNA modules and 9 classes were associated with risk overall at FDR ≤ 0.20. Triacylglycerols (TAG) showed heterogeneity by tumor aggressiveness (case-only heterogeneity P < 0.0001). The TAG association with risk overall and serous tumors differed by acyl carbon content and saturation. In summary, this study suggests that pseudouridine may be a novel risk factor for ovarian cancer and that TAGs may also be important, particularly for rapidly fatal tumors, with associations differing by structural features.

Significance: Pseudouridine represents a potential novel risk factor for ovarian cancer and triglycerides may be important particularly in rapidly fatal ovarian tumors.

Introduction

Ovarian cancer is the fifth leading cause of female cancer-related death in the United States (1). However, there are few known risk factors, such that current risk prediction models have a modest predictive capability, necessitating the identification of new risk factors to identify women at high risk.

Advances in technology have led to precise measures of small-molecule metabolites that are critical for growth and maintenance of cancer cells in biologic fluids (2). Several studies have identified metabolites as biomarkers of cancer risk. For example, branched chain amino acids were strongly associated with risk of pancreatic cancer (3) and lipid metabolites were inversely associated with risk of aggressive prostate cancer (4). Furthermore, prediagnostic serum concentrations of metabolites related to alcohol, vitamin E, and animal fats were modestly associated with ER+ breast cancer risk (5), while BMI-related metabolites were strongly related to increased risk (6). These findings support metabolomics profiling as a valuable strategy for identifying new cancer risk biomarkers.

Therefore, we used metabolomics assays to quantify several classes of circulating metabolites in plasma samples collected 3 to 23 years prior to ovarian cancer diagnosis within a nested case–control study, and, in an agnostic analysis, assessed their potential as biomarkers of ovarian cancer risk.

Materials and Methods

Study population

We conducted nested case–control studies within the Nurses Health Studies (NHS (7), NHSII (8)). The NHS was established in 1976 among 121,700 U.S. female nurses aged 30–55 years, and NHSII was established in 1989 among 116,429 female nurses aged 25–42 years. Participants have been followed biennially by questionnaire to update information on exposure status and disease diagnoses. Details are provided in the Supplementary File.

Incident cases of epithelial ovarian cancer were identified through biennial questionnaires or linkage with the National Death Index, for whom we obtained related medical records and pathology reports or linked to the relevant cancer registry when medical records were unattainable. A gynecologic pathologist reviewed the records to confirm the diagnosis and abstract date of diagnosis, invasiveness, stage, and histotype (serous, poorly differentiated (PD), endometrioid, clear cell (CC), mucinous, other/unknown), which is highly concordant
with centralized pathology review (9). Date of death was extracted from the death certificate.

Confirmed cases were diagnosed with ovarian cancer 3 years after blood collection until June 1, 2012 (NHS), or June 1, 2013 (NHSII); 252 cases of invasive and borderline epithelial ovarian cancer (212 in NHS and 40 in NHSII). We excluded cases diagnosed within 3 years of blood collection (N = 46) as most ovarian cancer cases are diagnosed at a late stage, with evidence suggesting preclinical disease up to 3 years before diagnosis (10). Cases were matched to one control on: cohort (NHS, NHSII); menopausal status and hormone therapy use at blood draw (premenopausal, postmenopausal/hormone therapy use, postmenopausal/no hormone therapy use, missing/unknown); menopausal status at diagnosis (premenopausal, postmenopausal, or unknown); age (± 1 year), date of blood collection (± 1 month); time of day of blood draw (± 2 hours); and fasting status (≥ 8 hours or ≤ 8 hours); women in NHSII who gave a lumpectomy sample were matched on the luteal date (date of the next period minus date of blood draw, ± 1 day).

Completion of the questionnaire was considered to imply informed consent when the study protocol was approved in 1976 (NHS) and 1989 (NHSII) by the institutional review boards of the Brigham and Women’s Hospital (Boston, MA) and Harvard T.H. Chan School of Public Health (Boston, MA), and those of participating registries as required. The studies were conducted in accordance with recognized ethical guidelines (Declaration of Helsinki).

Metabolite profiling

Plasma metabolites were profiled at the Broad Institute of MIT and Harvard (Cambridge, MA) using three complementary LC/MS-MS methods designed to measure polar metabolites and lipids as well as free fatty acids. Details are provided in the Supplementary File. In total, 608 known metabolites were measured. Metabolites with a coefficient of variation (CV) >25% or an intraclass correlation coefficient (ICC) <0.4 among blinded quality control (QC) samples were excluded (N = 132, Supplementary Table 1). Furthermore, metabolites with poor stability due to delayed processing (11) were excluded (N = 56, Supplementary Table 1). Included metabolites [e.g., amino acids, amino acids derivatives, amines, lipids, fatty acids, bile acids; N = 420 (69%)]; Supplementary Table 1] exhibited good reproducibility within person over 1 year (N11) and over 10 years. A total of 197 metabolites had no missing values among participant samples.

Statistical analysis

Identification of individual metabolites associated with risk

Missing values for metabolites (N = 211) with <10% missingness were imputed with 1/2 of the minimum value measured for that metabolite. We included a missing value indicator for metabolites (N = 12) with more than 10% missingness. Continuous metabolite values were transformed to probit scores to reduce the influence of skewed distributions and heavy tails on the results and to scale the measured metabolite values to the same range. Conditional logistic regression was used to evaluate metabolite associations, modeled continuously (with an additional indicator if >10% missingness), with risk of overall ovarian cancer. We present the ORs and 95% confidence intervals (95% CI) for an increase from the 10th to 90th percentile in metabolite levels or the indicator variable.

We compared conditional logistic regression to unconditional logistic regression adjusting for the matching factors and found similar results (Supplementary Table 1). Thus, subsequent analyses by histotype, rapidly fatal status, time between blood collection and diagnosis, and sensitivity analyses were conducted using the latter approach, allowing the use of all controls.

We conducted stratified analyses restricting to serous/PD tumors (cases = 176/controls = 252), endometrioid/CC tumors (cases = 34/controls = 252), rapidly fatal invasive cases (death occurring <3 years after diagnosis; cases = 86/controls = 252), and less aggressive invasive tumors (all other cases; cases = 138/controls = 252), as well as premenopausal (cases = 82/controls = 82) and postmenopausal women (cases = 137/controls = 137) at blood collection, and those diagnosed 3–11 years (cases = 121/controls = 252) and 12–23 years after blood collection (cases = 131/controls = 252). Models were adjusted for matching factors, duration of oral contraceptive use (none or <3 months, 3 months to 3 years, 3 to 5 years, >5 years), tubal ligation (yes/no), and parity (none, 1, 2, 3, 4 + children). We calculated heterogeneity by histotype, time to diagnosis, and tumor aggressiveness using case-only analyses and by menopausal status at blood collection by introducing an interaction term between the metabolite and menopausal status.

We conducted sensitivity analyses excluding borderline tumors (N = 25), known low-grade serous cases (N = 4), samples processed >24 hours after collection (N = 13 cases, N = 6 controls), cases with a diagnosis of a prior cancer (N = 17), or women with a diagnosis of another cancer after their matched case’s diagnosis (N = 35 controls). A permutation test (N = 5,000) was used to control the family-wise error rate (i.e., account for multiple testing) while accounting for the correlation structure of metabolites using the stepdown min P approach by Westfall and Young (12). Details are in the Supplementary File. We report unadjusted and multiple comparison adjusted P values and discuss individual metabolites associated with ovarian cancer risk at unadjusted P ≤ 0.01 given the hypothesis-generating nature of the study.

Identification of groups of metabolites associated with risk

Metabolite set enrichment analysis (MSEA; ref. 13) was used to identify groups of molecularly or biologically similar metabolites that were enriched among the metabolites associated with risk of overall ovarian cancer and histotypes and weighted gene coexpression network analysis (WGCNA; ref. 14) was used to identify metabolite modules and their association with ovarian cancer risk; details are in the Supplementary File. We report nominal P values and FDRs (15) for all metabolite groups and modules, discussing those at FDR ≤ 0.2. All analyses were performed using the statistical computing language R, version 3.5.0 (16).

Results

Study population

Of the 252 cases, 176 cases were diagnosed with serous/PD tumors, while 34 were classified as endometrioid/CC; 86 represented rapidly fatal tumors with death within 3 years of diagnosis (Table 1). Mean follow-up was 12.3 years. Distributions of ovarian cancer risk factors were generally in the expected directions.

Measured metabolites and their association with ovarian cancer risk

Of the 420 metabolites passing our QC filtering criteria, there were 159 lipids; 158 amino acids, amino acids derivatives amines, and cationic metabolites; and 103 free fatty acids, bile acids, and lipid mediators. Eight metabolites were associated with risk of
**Circulating Plasma Metabolites and Ovarian Cancer Risk**

**Table 1.** Characteristics of overall, serous/poorly differentiated and endometrioid-clear cell ovarian cancer cases, and all controls at time of blood collection.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>All controls (N = 252)</th>
<th>Overall OC (N = 252)</th>
<th>Serous/PD OC (N = 176)</th>
<th>Endometrioid/CC OC (N = 34)</th>
<th>Other histotypes (N = 42)</th>
<th>Rapidly fatal tumors (N = 86)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (SD)</td>
<td>55.6 (7.8)</td>
<td>55.5 (7.9)</td>
<td>55.3 (7.9)</td>
<td>54.0 (8.1)</td>
<td>57.8 (7.5)</td>
<td>58.5 (6.8)</td>
</tr>
<tr>
<td>Age at blood draw</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time to diagnosis (years)</td>
<td>51 (20)</td>
<td>49 (20)</td>
<td>50 (21)</td>
<td>50 (20)</td>
<td>55 (20)</td>
<td>53 (20)</td>
</tr>
<tr>
<td>BMI at blood draw</td>
<td>24.7 (4.1)</td>
<td>25.0 (4.7)</td>
<td>24.5 (4.2)</td>
<td>26.9 (5.8)</td>
<td>25.5 (5.4)</td>
<td>25.3 (5.3)</td>
</tr>
<tr>
<td>Oral contraceptive use duration</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parity</td>
<td>3 (1)</td>
<td>3 (1)</td>
<td>3 (1)</td>
<td>3 (1)</td>
<td>3 (1)</td>
<td>3 (1)</td>
</tr>
<tr>
<td>N (Percent)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tubal ligation</td>
<td>33 (13)</td>
<td>33 (13)</td>
<td>33 (13)</td>
<td>33 (13)</td>
<td>33 (13)</td>
<td>33 (13)</td>
</tr>
<tr>
<td>Menedoal status blood draw*</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Premenopausal, no HT use</td>
<td>82 (33)</td>
<td>82 (33)</td>
<td>56 (32)</td>
<td>16 (47)</td>
<td>10 (24)</td>
<td>14 (16)</td>
</tr>
<tr>
<td>Postmenopausal, HT use</td>
<td>71 (28)</td>
<td>68 (27)</td>
<td>47 (27)</td>
<td>5 (15)</td>
<td>16 (38)</td>
<td>29 (34)</td>
</tr>
<tr>
<td>Unknown</td>
<td>33 (13)</td>
<td>33 (13)</td>
<td>25 (14)</td>
<td>5 (15)</td>
<td>3 (7)</td>
<td>9 (10)</td>
</tr>
<tr>
<td>Cohort*</td>
<td>212 (84)</td>
<td>212 (84)</td>
<td>147 (84)</td>
<td>27 (79)</td>
<td>38 (90)</td>
<td>79 (92)</td>
</tr>
<tr>
<td>Race</td>
<td>251 (100)</td>
<td>251 (100)</td>
<td>175 (100)</td>
<td>34 (100)</td>
<td>42 (100)</td>
<td>86 (100)</td>
</tr>
<tr>
<td>Oral contraceptive use duration</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>None or &lt;3 months</td>
<td>123 (49)</td>
<td>118 (47)</td>
<td>81 (46)</td>
<td>18 (53)</td>
<td>19 (45)</td>
<td>48 (56)</td>
</tr>
<tr>
<td>3 months to 3 years</td>
<td>33 (13)</td>
<td>33 (13)</td>
<td>22 (12)</td>
<td>3 (9)</td>
<td>7 (17)</td>
<td>10 (12)</td>
</tr>
<tr>
<td>5 to 5 years</td>
<td>45 (18)</td>
<td>63 (25)</td>
<td>46 (26)</td>
<td>8 (24)</td>
<td>9 (21)</td>
<td>16 (19)</td>
</tr>
<tr>
<td>5+ years</td>
<td>51 (20)</td>
<td>39 (15)</td>
<td>27 (15)</td>
<td>5 (15)</td>
<td>7 (17)</td>
<td>12 (14)</td>
</tr>
<tr>
<td>Parity</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No children</td>
<td>12 (5)</td>
<td>24 (10)</td>
<td>16 (9)</td>
<td>5 (15)</td>
<td>3 (7)</td>
<td>7 (8)</td>
</tr>
<tr>
<td>1 child</td>
<td>11 (4)</td>
<td>13 (5)</td>
<td>8 (5)</td>
<td>1 (3)</td>
<td>4 (10)</td>
<td>2 (2)</td>
</tr>
<tr>
<td>2 children</td>
<td>72 (29)</td>
<td>89 (35)</td>
<td>60 (34)</td>
<td>15 (44)</td>
<td>14 (33)</td>
<td>23 (27)</td>
</tr>
<tr>
<td>3 children</td>
<td>77 (31)</td>
<td>65 (26)</td>
<td>45 (26)</td>
<td>9 (26)</td>
<td>11 (26)</td>
<td>25 (29)</td>
</tr>
<tr>
<td>4+ children</td>
<td>80 (32)</td>
<td>61 (24)</td>
<td>47 (27)</td>
<td>4 (12)</td>
<td>10 (24)</td>
<td>29 (34)</td>
</tr>
<tr>
<td>Risk status</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race</td>
<td>212 (84)</td>
<td>212 (84)</td>
<td>147 (84)</td>
<td>27 (79)</td>
<td>38 (90)</td>
<td>79 (92)</td>
</tr>
<tr>
<td>Serum/plasma</td>
<td>30 (12)</td>
<td>30 (12)</td>
<td>30 (12)</td>
<td>30 (12)</td>
<td>30 (12)</td>
<td>30 (12)</td>
</tr>
<tr>
<td>Metabolites</td>
<td>212 (84)</td>
<td>212 (84)</td>
<td>147 (84)</td>
<td>27 (79)</td>
<td>38 (90)</td>
<td>79 (92)</td>
</tr>
<tr>
<td>Abbreviation: HT, any type of hormone therapy; OC, ovarian cancer.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*Matching factors.</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Overall ovarian cancer at a nominal P ≤ 0.01 (Table 2A; Fig. 1; Supplementary Table 1). ORs for an increase from the 10th to the 90th percentile of levels ranged between 0.49 and 2.56. The top three metabolites associated with risk were pseudouridine (OR = 2.56; 95% CI, 1.48–4.45; P = 0.001), C18:0 sphingomyelin (SM; OR = 2.10; 95% CI, 1.26–3.49; P = 0.004), and 4-acetamidobutanoate (OR = 2.10; 95% CI, 1.24–3.56; P = 0.006). Pseudouridine had an adjusted P = 0.15 (accounting for all tested metabolites and their correlation structure); all other metabolites had adjusted P > 0.5. The test of the global null hypothesis that no metabolite was associated with risk had P = 0.15. Results did not change in sensitivity analyses excluding specific case and control populations (Supplementary Tables 1.2–1.5).

Five metabolites were associated with risk of serous/PD tumors at a nominal P ≤ 0.01 (Table 2B; Fig. 1; Supplementary Table 2). ORs for an increase from the 10th to the 90th percentile of metabolites levels for these metabolites ranged between 1.99 and 2.38. The top three metabolites were pseudouridine (OR = 2.38; 95% CI, 1.33–4.32; P = 0.004), C52:5 triacylglycerol (TAG; OR = 2.09; 95% CI, 1.23–3.59; P = 0.007), and C52:4 TAG (OR = 2.03; 95% CI, 1.21–3.47; P = 0.008). However, none of the metabolites remained significant after accounting for multiple comparisons via permutation (adjusted P > 0.55). The test of the global null hypothesis that no metabolite was associated with risk had P = 0.53. Results did not change in sensitivity analyses in which we excluded low-grade serous cases (Supplementary Table 2.1).

Thirty metabolites were associated with risk of endometrioid/CC tumors at a nominal P ≤ 0.01 (Table 2C; Fig. 1; Supplementary Table 2). ORs for an increase from 10th to the 90th percentile of metabolites levels for these metabolites ranged between 0.11 and 0.24 for inverse associations, and between 3.85 and 9.84 for positive associations. The top three metabolites positively associated with risk were pseudouridine (OR = 9.84; 95% CI, 2.89–37.82; P = 0.0003), C2 carnitine (OR = 7.4; 95% CI, 2.37–25.35; P = 0.001), and C56:7 TAG (OR = 5.85; 95% CI, 2.04–18.02; P = 0.001). The top three metabolites inversely associated with risk were C36:2 phosphatidylcholines (PC) plasmalogen (OR = 0.11; 95% CI, 0.03–0.35; P = 0.0003), C34:1 PC plasmalogen-A (OR = 0.18; 95% CI, 0.05–0.54; P = 0.003), and C22:0 lysophosphatidylethanolamine (LPE; OR = 0.21; 95% CI, 0.07–0.59; P = 0.004). C36:2 PC plasmalogen and pseudouridine had an adjusted P = 0.06 and 0.07, respectively (accounting for all tested metabolites and their correlation structure). All other metabolites had adjusted P ≥ 0.14. The test of the global null hypothesis that no metabolite was associated with risk had P = 0.06.
On the individual metabolite level, histograms and QQ-plots of the nominal P values (Supplementary Fig. 1) together with the results of the permutation-based approach to account for testing multiple correlated metabolites (Westfall and Young’s stepdown min P approach) suggest the existence of a metabolomic signal for overall ovarian cancer and nonserous tumors.

**Metabolite groups associated with risk of ovarian cancer**

In the MSEA analysis, nine metabolite groups were enriched among metabolites associated with risk of ovarian cancer overall at an FDR ≤ 0.2 (Fig. 2; Supplementary Table 3). The top five groups were organic acids and derivatives, PE plasmalogens, TAGs, cholesteryl esters, and PC plasmalogens. Nine metabolite groups were associated with risk of serous/PD tumors with FDR ≤ 0.20 (Fig. 2; Supplementary Table 3). The top five were as follows: nucleosides, nucleotides, and analogues; PC plasmalogens; carnitines; sphingomyelins; and PE plasmalogens. Finally, eleven metabolite groups were associated with risk of endometrioid/CC tumors at FDR ≤ 0.20 (Fig. 2; Supplementary Table 3). The top five associated metabolite groups were TAGs, DAGs, fatty acyls, lysophosphatidylserines (LPS), and carnitines. TAGs were enriched in the above at FDR ≤ 0.05. Notably, we observed differential associations by acyl carbon number and double bond content with risk of ovarian cancer overall (Supplementary Fig. 2) and serous/PD tumors (Supplementary Fig. 3), but not with endometrioid/CC tumors (Supplementary Fig. 4). Specifically, TAGs with higher number of acyl carbon atoms and double bonds were associated with increased risk, while TAGs with lower number of acyl carbon atoms and double bonds were associated with decreased risk. We did not observe similar patterns for other lipid classes (Supplementary Fig. 2–4).

**Metabolite modules associated with risk of ovarian cancer**

WGCNA identified seven metabolite modules associated with risk of ovarian cancer with FDR ≤ 0.20 (Table 3 and Fig. 3A–D). Module 1 [M1, characterized by steroids and steroid derivatives, organic acids and derivatives, and organonitrogen compounds (Supplementary Fig. 5–7; Supplementary Table 4)], M2 (characterized by TAGs, PCs, PE, LPCs, and LPEs), M6 (characterized by TAGs, LPEs, and CEs), and M7 (characterized by TAGs, DAGs, ceramides, and CEs) were associated with increased risk of ovarian cancer overall, OR, increase from 10th to 90th percentile = 1.99 (P = 0.013/FDR = 0.072), 1.62 (P = 0.093/FDR = 0.186), 1.56 (P = 0.081/FDR = 0.186), and 1.8 (P = 0.015/FDR = 0.072), respectively. M4 [characterized by carnitines, pseudouridine...
(inversely weighted), and organic acids and derivatives was associated with decreased risk (OR = 0.5; P = 0.022/FDR = 0.072). M7 was associated with increased risk of serous/PD tumors (OR = 1.97; P = 0.012/FDR = 0.117). Finally, four modules were associated with risk of endometrioid/CC tumors: M2 (OR = 6.14; P = 0.002/FDR = 0.011), M4 (OR = 0.17; P = 0.003/FDR = 0.011), M5 (PC and PE plasmalogens; OR = 0.35; P = 0.041/FDR = 0.072), and M8 (fatty acyls, OR = 0.22; P = 0.007/FDR = 0.017).

Metabolites associated with ovarian cancer risk by menopausal status at blood collection

C22:0 LPS isomer was suggestively associated with increased risk among postmenopausal women (OR = 1.83; 95% CI, 0.92–3.63; P = 0.085) and decreased risk among premenopausal women (OR = 0.44; 95% CI, 0.17–1.08; P = 0.074), with a heterogeneity P = 0.004 (Supplementary Table 5). C38:4 PC plasmalogens was suggestively associated with increased risk among postmenopausal women (OR = 1.92; 95% CI, 0.96–3.85; P = 0.066) and decreased risk among premenopausal women (OR = 0.16; 95% CI, 0.05–0.51; P = 0.002), with a heterogeneity P = 0.005. Among premenopausal women, 14/22 (63%) metabolites associated with risk at P ≤ 0.1 were inversely related, but among premenopausal women only 15/98 (15%) metabolites showed inverse associations. Pseudouridine did not show heterogeneity by menopausal status (heterogeneity P = 0.32).

Metabolites associated with ovarian cancer risk by time between blood collection and diagnosis

Hydroxyvitamin D3 was associated with increased risk among participants with blood collection 12–23 years before diagnosis (OR = 1.84; 95% CI, 1.02–3.37; P = 0.044) but not among participants with blood collection 3–11 years before diagnosis (OR = 0.64; 95% CI, 0.36–1.15; P = 0.141), with a heterogeneity P = 0.002 (Supplementary Table 6). C40:6 phosphatidylserine (PS) was associated with decreased risk among participants with blood collection 12–23 years before diagnosis (OR = 0.55; 95% CI, 0.30–0.99; P = 0.049) but not among participants with blood collection 3–11 years before diagnosis (OR = 1.41; 95% CI, 0.79–2.51; P = 0.245), with a heterogeneity P = 0.008. Pseudouridine showed suggestively stronger associations (heterogeneity P = 0.066) among women for whom sample collection was 3–11 years before diagnosis (OR = 4.48; 95% CI, 2.25–9.24;
Metabolites associated with ovarian cancer risk by tumor aggressiveness

Fifty-three lipid-related metabolites (26 TAGs, 7 PCs, 6 LPEs, 3 PEs, 3 LPC, 4DAGs, 2 LPSs, and 2 PSs) showed differences by tumor aggressiveness at heterogeneity \( P \leq 0.01 \) (Supplementary Table 7). Seven metabolites (6 TAGs and 1 PS) were associated with increased risk of rapidly fatal disease with ORs ranging between 2.56 and 3.07 at \( P \leq 0.008 \), but not with less aggressive tumors (\( P > 0.62 \)), with heterogeneity \( P \leq 0.001 \). Several lipid-related metabolite classes (DAGs, LPCs, LPEs, PEs, PSs, and TAGs with high acyl carbon content and saturation) were overrepresented in rapidly fatal tumors versus controls, while carnitines were overrepresented in less aggressive tumors (Supplementary Fig. 8, panels A and B). TAGs with lower acyl carbon content and saturation were inversely associated with less aggressive tumors. Pseudouridine did not show heterogeneity by tumor aggressiveness (heterogeneity \( P = 0.13 \)).

Discussion

We conducted the first large-scale agnostic analysis of metabolomics and risk of ovarian cancer. We identified a potential novel risk factor, plasma pseudouridine, which was associated with an increased risk of ovarian cancer overall and nonserous tumors and suggestively for serous/PD disease. Stronger associations for pseudouridine were observed among cases diagnosed within 3–11 years after blood collection. We identified several metabolite groups and metabolite modules associated with risk of ovarian cancer risk, as well as multiple subtype-specific associations, that open up new opportunities for assessing novel metabolite pathways involved in ovarian cancer development.

Pseudouridine

Pseudouridine is the most abundant posttranscriptionally modified nucleoside and is an isomer of uridine. It is produced by pseudouridine synthase by isomerizing uridines from transfer RNA, which is involved in protein translation or spliceosomal snRNA, which plays a role in pre-mRNA splicing. Pseudouridine was nominally associated with
Table 3. \( P \) values, FDR, and odds ratios for an increase from the 10th to the 90th percentile of metabolite levels and 95% confidence intervals of WGCNA metabolite modules associated with risk of ovarian cancer overall and by histotype.

<table>
<thead>
<tr>
<th>Module/number</th>
<th>Explained variance [%]</th>
<th>OR (95% CI)</th>
<th>P</th>
<th>Odds ratio</th>
<th>95% CI</th>
<th>FDR</th>
<th>P</th>
<th>Explained variance [%]</th>
<th>OR (95% CI)</th>
<th>P</th>
<th>Odds ratio</th>
<th>95% CI</th>
<th>FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1/24</td>
<td>1.99 (1.15–3.42)</td>
<td>0.013</td>
<td>0.072</td>
<td>0.013</td>
<td>0.072</td>
<td>0.013</td>
<td>0.072</td>
<td>0.013</td>
<td>0.072</td>
<td>0.013</td>
<td>0.072</td>
<td>0.013</td>
<td>0.072</td>
</tr>
<tr>
<td>M2/79</td>
<td>2.88 (1.49–5.61)</td>
<td>0.092</td>
<td>0.494</td>
<td>0.092</td>
<td>0.494</td>
<td>0.092</td>
<td>0.494</td>
<td>0.092</td>
<td>0.494</td>
<td>0.092</td>
<td>0.494</td>
<td>0.092</td>
<td>0.494</td>
</tr>
<tr>
<td>M3/76</td>
<td>2.74 (1.62–4.66)</td>
<td>0.103</td>
<td>0.397</td>
<td>0.103</td>
<td>0.397</td>
<td>0.103</td>
<td>0.397</td>
<td>0.103</td>
<td>0.397</td>
<td>0.103</td>
<td>0.397</td>
<td>0.103</td>
<td>0.397</td>
</tr>
<tr>
<td>M4/74</td>
<td>2.69 (1.62–4.53)</td>
<td>0.104</td>
<td>0.398</td>
<td>0.104</td>
<td>0.398</td>
<td>0.104</td>
<td>0.398</td>
<td>0.104</td>
<td>0.398</td>
<td>0.104</td>
<td>0.398</td>
<td>0.104</td>
<td>0.398</td>
</tr>
<tr>
<td>M5/79</td>
<td>2.63 (1.61–4.47)</td>
<td>0.105</td>
<td>0.399</td>
<td>0.105</td>
<td>0.399</td>
<td>0.105</td>
<td>0.399</td>
<td>0.105</td>
<td>0.399</td>
<td>0.105</td>
<td>0.399</td>
<td>0.105</td>
<td>0.399</td>
</tr>
<tr>
<td>M6/74</td>
<td>2.58 (1.59–4.27)</td>
<td>0.106</td>
<td>0.400</td>
<td>0.106</td>
<td>0.400</td>
<td>0.106</td>
<td>0.400</td>
<td>0.106</td>
<td>0.400</td>
<td>0.106</td>
<td>0.400</td>
<td>0.106</td>
<td>0.400</td>
</tr>
<tr>
<td>M7/79</td>
<td>2.56 (1.58–4.24)</td>
<td>0.107</td>
<td>0.401</td>
<td>0.107</td>
<td>0.401</td>
<td>0.107</td>
<td>0.401</td>
<td>0.107</td>
<td>0.401</td>
<td>0.107</td>
<td>0.401</td>
<td>0.107</td>
<td>0.401</td>
</tr>
<tr>
<td>M8/24</td>
<td>2.54 (1.56–4.16)</td>
<td>0.108</td>
<td>0.402</td>
<td>0.108</td>
<td>0.402</td>
<td>0.108</td>
<td>0.402</td>
<td>0.108</td>
<td>0.402</td>
<td>0.108</td>
<td>0.402</td>
<td>0.108</td>
<td>0.402</td>
</tr>
<tr>
<td>M9/14</td>
<td>2.53 (1.55–4.03)</td>
<td>0.109</td>
<td>0.403</td>
<td>0.109</td>
<td>0.403</td>
<td>0.109</td>
<td>0.403</td>
<td>0.109</td>
<td>0.403</td>
<td>0.109</td>
<td>0.403</td>
<td>0.109</td>
<td>0.403</td>
</tr>
<tr>
<td>M10/11</td>
<td>2.52 (1.54–4.01)</td>
<td>0.110</td>
<td>0.404</td>
<td>0.110</td>
<td>0.404</td>
<td>0.110</td>
<td>0.404</td>
<td>0.110</td>
<td>0.404</td>
<td>0.110</td>
<td>0.404</td>
<td>0.110</td>
<td>0.404</td>
</tr>
</tbody>
</table>

Circulating Plasma Metabolites and Ovarian Cancer Risk

risk overall and for both histotypes, with no significant heterogeneity (\( P = 0.16 \)) by histotype. This suggests that pseudouridine may represent a common etiologic mechanism underlying different histotypes of ovarian cancer, which has been observed for other risk factors, such as aspirin and C-reactive protein. In retrospective studies, pseudouridine was elevated in urine (17) and plasma (18) from patients with epithelial ovarian cancer versus healthy controls. This, in combination with our finding that pseudouridine had a stronger association when assessed 3–11 years before diagnosis, suggests that this modified nucleotide may be important in progression of preclinical lesions to fully overt invasive disease, which for high-grade serous ovarian cancer appears to be about 7–9 years (19). Increasing evidence suggests that pseudouridylation plays a role in cancer-associated splicing distributions, which are more variable than in normal tissues. Notably, tissue-specific alternative splicing reverts to a default cancer pattern that directly contributes to cellular transformation and cancer progression (20). This has been observed in serous carcinomas, which have highly dysregulated splicing compared with normal tissue (21). Furthermore, aberrant pseudouridylation may lead to altered and reduced translational fidelity of p53 (22), which is mutated in nearly all high-grade serous tumors (23). Another potential mechanism is via circular RNA activity, which is altered because of isomerization of uridine to pseudouridine, and has been shown to be dysregulated in ovarian cancer (24). Interestingly, pseudouridine is associated with the estimated glomerular filtration rate, a marker of kidney function (25). While kidney function alters the immune response, potentially contributing to the development of cancer (26), associations with cancer incidence are mixed, although a recent study observed that thiazide diuretics, which can affect kidney function, are associated with a higher risk of ovarian cancer (27). Additional research should explore the potential role of pseudouridine in precursor lesions to ovarian cancer, the relation between circulating pseudouridine to ovarian and fallopian tube tissue levels, and if kidney function plays a role in the initiation or development of ovarian cancer.

**Triacylglycerides**

Notably, several individual TAGs were nominally related to risk and showed significantly stronger associations with rapidly fatal tumors. Evidence suggests that established ovarian cancer risk factors vary by tumor aggressiveness (28). As high-grade serous ovarian cancer was the predominant histotype among rapidly fatal as well as less aggressive tumors, our data suggest that there are potential differences between the metabolic profiles of these two groups of tumors independent of histotype. TAGs as a group were enriched in the MSEA analysis, and 3 of 7 WCGNA modules related to risk were characterized by TAGs. Long chain fatty acids, a main source of energy in the human body, are stored and transported from the small intestine and liver to peripheral cells as TAGs (29). Lipid synthesis and metabolism that release free fatty acids from TAGs are dysregulated in ovarian tumors, increasing cell migration and invasive potential (30). Furthermore, several human studies reported suggestive associations of ovarian cancer risk with total cholesterol (positive; ref. 31) or high-density lipoprotein (inverse; ref. 32). In addition, ovarian cancer metastasizes preferentially to the adipose-rich omentum (33). Omental fat possesses a distinct lipidomic signature with several lipid groups, including TAGs, DAGs, and SMs, showing differences when compared with subcutaneous fat (34). Finally, plasma TAGs represent known risk factors for cardiovascular disease and coronary heart disease. A recent study identified that TAGs at the extremes of carbon atom and saturation had differential associations with diabetes risk (35). We also observed differential associations by TAG fatty
acids length and saturation, with higher number of carbon atoms and double bonds related to an increased risk and lower number of carbon atoms and double bonds related to decreased risk, particularly for serous/PD tumors. A similar pattern was observed in a retrospective study of serum samples from high-grade serous ovarian cancer cases and controls (36). Together with our results, these findings suggest that circulating TAG levels may be a risk biomarker for ovarian cancer, particularly for rapidly fatal tumors. Additional prospective studies are needed to validate these associations in different populations and assess the potential differential role of various TAG species in ovarian carcinogenesis.

Other metabolite groups

A number of metabolite groups and classes were associated with ovarian cancer risk, including organic acids and derivatives, and SMs, the latter of which was hypothesized a priori as a potential risk biomarker and is discussed elsewhere (37). A metabolite module driven by carnitines, organic acids and derivatives, and carboxylic acids and derivatives, which included pseudouridine (highly negatively weighted), was associated with decreased risk of overall ovarian cancer and nonserous tumors. This module includes asymmetric dimethylarginine, which has been related to risk of cardiovascular disease (38), and inhibits nitric oxide synthesis and may have antiproliferative properties (39) including in ovarian tumors (40). LPEs were also represented in WCGNA modules associated with increased risk of overall and endometrioid/CC ovarian cancers. LPEs have been shown to increase migration in response to chemotherapy as well as have invasive potential in ovarian cancer cell lines (41). In MSEA analyses, several metabolite classes had a significant negative enrichment score, including PE plasmalogens, PC plasmalogens, and cholesteryl esters, independent of subtype.

Little work has examined these markers in ovarian cancer development or etiology. Sphingolipids [(SL); including SMs, PCs, PEs, LPCs, LPEs, cholesteryl esters, acylcarnitines] are associated with a series of conditions that may be related to ovarian cancer, including thrombosis in a mouse study (42), myocardial infarction among patients with symptomatic coronary artery disease (43), type I and type II diabetes in human studies (44), diabetic kidney disease in...
mice (45, 46), and airway inflammation and asthma in mouse and human studies (47–49). Notably, patients with ovarian cancer have the highest incidence of venous thromboembolism of all solid tumor types (50), which is significantly related to higher mortality (51). Coagulation activation by tumors promotes development of venous thromboembolism, which in turn favors cancer progression through tumor growth, angiogenesis, invasion, immune eva-
sion, and metastasis (52). Furthermore, cholesterol-lowering statins have anti-inflammatory, antiproliferative, apoptotic, and anti-
vasive qualities (53–57), and can lower SMs (58). Data on ovarian cancer risk reduction by statins have been mixed. A meta-analysis of existing studies suggested a lower risk for ovarian cancer associated with statin use (59) while postdiagnostic statin use was inversely associated with overall survival and ovarian cancer–specific mortality (60–62). Additional work should evaluate whether these conditions and medications associated with the identified metabolites represent novel risk factors for ovarian cancer, preferably using large consortia to ensure power.

Our study has several strengths and limitations. Importantly, this is a prospective study of ovarian cancer risk with coverage of multiple different metabolite classes. Additional strengths include the long follow-up time and detailed covariate information. Our cohort consisted of registered nurses, a group that are not representative of the general population (e.g., social economic status), although established risk factor associations in these cohorts are similar to those in other more representative studies (63). While we had over 250 ovarian cancer cases and controls, we had more limited sample sizes for specific histotypes, which have been shown to have different associations for known risk factors (64). We used medical records and pathology reports to confirm diagnosis and extract histotype and cannot rule out the possibility of histotype misclassification, although we previously showed high concordance to centralized slide review (9). To maximize power, borderline and tumors of unknown morphology were included. We did not include information on family history of ovarian cancer. However, only 2 cases were diagnosed before age 45, suggesting that early onset disease, likely due to high-risk mutations, does not play a role in this study. We also applied stringent QC criteria to limit identification of spurious associations. Another limitation is that we only analyzed blood samples collected at one point in time; however, we demonstrate that the majority of the measured metabolites have a high within person stability over time (11). Furthermore, we do not have an independent validation dataset. As this type of data becomes more common, further population studies are needed to validate the results discussed here, while experimental studies are required to understand the biological mechanisms underlying these associations.

In summary, circulating levels of plasma pseudouridine were associated with higher risk of ovarian cancer 3–23 years before diagnosis, with stronger associations among participants with samples collected closer to diagnosis. In addition, several metabolite groups and metabolite modules were associated with risk of disease overall and by subtype. While independent prospective studies are needed for validation, our results highlight some potentially important novel metabolites that may play a role in the etiology of ovarian cancer. Potential experimental studies to understand the biological mechanisms of these risk biomarkers could examine their role in carcinogenic tendencies in ovarian and fallopian tube cancer cell lines (with and without p53 mutations), mouse models of early and late ovarian lesions leading to ovarian cancer, as well as xenograft mouse models in which pseudouridine production has been impaired or enhanced. Adding these new risk biomarkers to current risk prediction models may help the identification of high-risk women.

Disclosure of Potential Conflicts of Interest
No potential conflicts of interest were disclosed.

Disclaimer
The authors assume full responsibility for analyses and interpretation of these data.

Authors' Contributions
Conception and design: P. Kraft, E.M. Poole, S.S. Tworoger
Development of methodology: O.A. Zeleznik
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): A.H. Elaissen, E.M. Poole, S. Jeanfave, A.A. Deik, K. Bullock, D.S. Hitchcock, J. Avila-Pacheco, C.B. Clish, S.S. Tworoger
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): O.A. Zeleznik, P. Kraft, E.M. Poole, B.A. Rosner, A.A. Deik, J. Avila-Pacheco, S.S. Tworoger
Writing, review, and/or revision of the manuscript: O.A. Zeleznik, A.H. Elaissen, P. Kraft, E.M. Poole, D.S. Hitchcock, J. Avila-Pacheco, C.B. Clish, S.S. Tworoger
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): O.A. Zeleznik, A.H. Elaissen, E.M. Poole
Study supervision: S.S. Tworoger

Acknowledgments
The authors would like to thank the participants and staff of the Nurses Health Studies and Nurses Health Studies II for their valuable contributions as well as the following state cancer registries for their help: AL, AZ, AR, CA, CO, CT, DE, FL, GA, ID, IL, IN, IA, KY, LA, ME, MD, MA, MI, NE, NH, NJ, NY, NC, ND, OH, OK, OR, PA, RI, SC, TN, TX, VA, WA, WY. This study was funded by the NCI (F01 CA087969, U01 CA176726, UM1 CA186107) and the Department of Defense (W81XWH-12-1-0561).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received August 19, 2019; revised December 16, 2019; accepted January 17, 2020, published first January 22, 2020.

References
Zelevnik et al.


60. Lavie O, Pinchev M, Rennert HS, Segev Y, Rennert G. The effect of statins on risk and survival of gynecological malignancies. Gynecol Oncol 2013;130:615–9.
A Prospective Analysis of Circulating Plasma Metabolites Associated with Ovarian Cancer Risk

Oana A. Zeleznik, A. Heather Eliassen, Peter Kraft, et al.


Updated version
Access the most recent version of this article at:
doi:10.1158/0008-5472.CAN-19-2567

Supplementary Material
Access the most recent supplemental material at:
http://cancerres.aacrjournals.org/content/suppl/2020/01/22/0008-5472.CAN-19-2567.DC1

Cited articles
This article cites 60 articles, 11 of which you can access for free at:
http://cancerres.aacrjournals.org/content/80/6/1357.full#ref-list-1

Citing articles
This article has been cited by 2 HighWire-hosted articles. Access the articles at:
http://cancerres.aacrjournals.org/content/80/6/1357.full#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, use this link
http://cancerres.aacrjournals.org/content/80/6/1357.
Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.