UNIVERSITAT DE
BARCELONA

Dietary intake and biomarkers
of acrylamide exposure and risk
# of endometrial and ovarian cancer

A molecular epidemiologic study in the European
Prospective Investigation into Cancer and Nutrition

Mireia Obón Santacana

RATIONALE

2. RATIONALE

As shown in the previous chapter, acrylamide is classified by IARC as a Group 2A *probably carcinogenic to humans*, and it is found in relatively high concentrations in typically consumed foods (i.e., chips, crispbread, cereals, and coffee), cigarette smoke, and in various industries. Although acrylamide exerts different effects in the body, we were interested in study the link between dietary acrylamide consumption and cancer risk, since acrylamide exposure has been clearly associated to carcinogenicity in multiple tissues in animal models.

Several epidemiologic studies evaluating dietary acrylamide exposure (based on FFQs) and cancer risk (i.e., bladder, brain, breast, colon, colorectal, endometrial, esophageal, kidney, ovarian, laryngeal, prostate, stomach, and rectum) were published after the Swedish researchers discovery of acrylamide in food. No convincing evidence from these studies was found; however, two studies (one prospective study from the Netherlands and one nested case-control from Denmark) observed weak evidence for positive postmenopausal breast cancer risk with acrylamide exposure. Moreover, two prospective cohort studies (one from the Netherlands and one from the US) observed that women who were classified at the highest quintiles of acrylamide intake had higher risk of developing EC and EOC than women classified in the lowest intakes. One case-control study from Italy and one prospective cohort study from Sweden did not report associations with these cancers.

Thus, there was very limited research from large prospective cohort studies on the association between dietary intake of acrylamide EC or EOC risk. Using data from the EPIC cohort study, we will have the chance to determine if women who have higher dietary acrylamide exposure are at a higher risk of EC and EOC, after adjusting for confounding variables. The EPIC study is one of the largest prospective cohort studies in nutrition and cancer, and this large sample size will allow us to examine further associations by histological EC and EOC subtypes, that could not be deeply investigated in previous studies.

There were no published studies that evaluated the association between acrylamide and glycidamide blood levels and the risk of developing EC or EOC before the start of this thesis (October 2012). Besides epidemiological data, blood samples were also collected at recruitment (before cancer diagnosis) in 80% of the EPIC population. We will use red blood cells to measure HbAA and HbGA levels in women. Biomarkers of acrylamide exposure are valid measures of human internal dose, and avoid possible measurement errors from using acrylamide intake estimates based on FFQs. We will investigate if diet and lifestyle variables determine biomarkers levels. Finally, we will evaluate if women with higher biomarkers levels in the blood are at a higher risk of EC or EOC. The first study that investigated the link between acrylamide biomarkers and EOC risk was published in 2013.

The work of this thesis will contribute to the existing scientific evidence to determine if acrylamide exposure appears to play a role in endometrial and/or ovarian carcinogenesis.

HYPOTHESIS AND AIMS

3. HYPOTHESIS AND AIMS

The general aim of this thesis is to assess the risk of endometrial and epithelial ovarian cancer among women who consume acrylamide through their diet. The following hypothesis and aims described below correspond to the five published papers presented in this thesis.

__Hypothesis 1__: High dietary acrylamide intake is associated with increased risk of endometrial cancer. Risk differs by tobacco smoking, alcohol consumption and oral contraceptive use.

- **Objective 1.1:** Estimate adjusted relative risk for high vs. low acrylamide intake and overall endometrial cancer (1,382 cases) and type-I endometrial cancer (627 cases).
- **Objective 1.2:** Describe acrylamide intake and cancer associations by smoking status, geographic region, alcohol consumption, body mass index (BMI), hormone-replacement therapy (HRT), and oral contraceptive use (OC).

__Hypothesis 2__: High dietary acrylamide intake is associated with increased risk of epithelial ovarian cancer. Risk differs by tobacco smoking, alcohol consumption and oral contraceptive use.

- **Objective 2.1:** Estimate adjusted relative risk for high vs. low acrylamide intake and epithelial ovarian cancer (1,191 cases), and between epithelial ovarian histologic subtypes.
- **Objective 2.2:** Describe acrylamide intake and cancer associations by smoking status, geographic region, alcohol consumption, BMI, HRT, and OC use.

__Hypothesis 3__: Specific foods and lifestyle variables determine biomarker levels of acrylamide and its metabolite, glycidamide in non-smoking postmenopausal women. Biomarker levels are modified by heavy alcohol consumption.

- **Objective 3.1:** Measure pre-diagnostic hemoglobin adduct levels of acrylamide (HbAA) and its metabolite, glycidamide (HbGA), in red blood cells from the nested case-control studies in the EPIC cohort.
- **Objective 3.2:** Identify determinants of HbAA and HbGA adduct levels, their sum (HbAA+HbGA) and their ratio (HbGA/HbAA) in 801 control women using information on known sources of dietary acrylamide exposure and lifestyle variables.
- **Objective 3.3:** Evaluate effect measure modification by baseline alcohol intake, BMI, and geographic region.

__Hypothesis 4__: High biomarker levels of HbAA and HbGA are associated with increased endometrial cancer risk. Heavy alcohol consumption and oral contraceptive use modify effects.

- **Objective 4.1:** Estimate adjusted relative risks of overall endometrial cancer (383 cases and 385 controls), and type-I endometrial cancer (171 cases) according to biomarker levels of HbAA and HbGA.
- **Objective 4.2:** Explore modification of biomarker–cancer associations by alcohol consumption, BMI, HRT, and OC use.

Hypothesis 5: _High biomarker levels of HbAA and HbGA are associated with increased epithelial ovarian cancer risk. Heavy alcohol consumption and oral contraceptive use modify effects._

- **Objective 5.1:** Estimate adjusted relative risks of epithelial ovarian cancer (334 cases and 417 controls) and tumor histologic subtypes according to biomarker levels of HbAA and HbGA.
- **Objective 5.2:** Explore modification of biomarker–cancer associations by alcohol consumption, BMI, HRT, and OC use.

MATERIALS AND METHODS

4. MATERIALS AND METHODS

The present chapter summarizes the materials and methods described in detail in all five papers that comprise this thesis.

4.1 The EPIC study

The EPIC study is defined as an ongoing multicenter prospective cohort study which comprises twenty three research centers in ten European countries including Denmark (Aarhus, Copenhagen), France, Germany (Heidelberg, Potsdam), Greece, Italy (Florence, Turin, Varese, Naples, Ragusa), The Netherlands (Bilthoven, Utrecht), Norway, Spain (Granada, Murcia, Asturias, Navarra, San Sebastian), Sweden (Malmö, Umeå), and the United Kingdom (Norfolk, Oxford) (**Figure 10**).
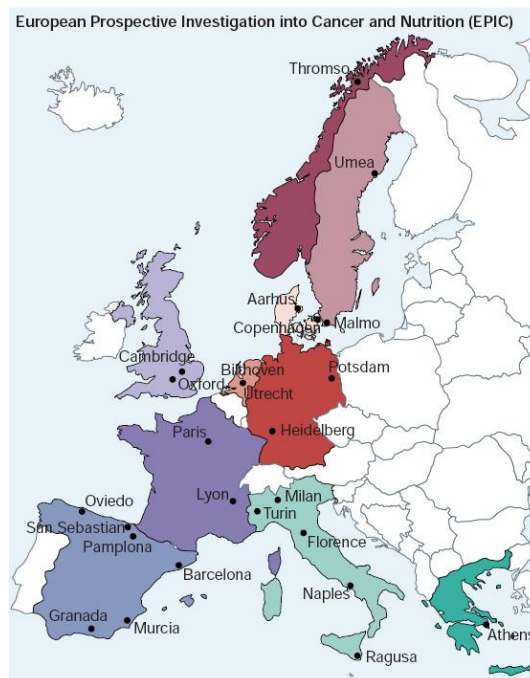
Figure 10. EPIC collaborating centers [126]

The EPIC study originally included 519,978 participants, of which 366,521 were women. Recruitment started between 1992 and 1998 at ages 35 to 70 years. Participants were mostly recruited from the general population; however, there were a few exceptions: a) The French cohort was based on members of the health insurance for teachers. b) Components of the Italian and the Spanish cohorts included members of local blood donors associations. c) The Utrecht (The Netherlands) and the Florence (Italy) cohorts included women invited from a local population-based breast cancer screening program. d) Half of the Oxford (United Kingdom) cohort was recruited among subjects who did not consume meat (including vegans, lacto-ovo-vegetarians, and fish consumers). e) Cohorts from France, Norway, Utrecht, and Naples (Italy) only recruited women [127].

4.2 Identification of cancer cases

Information on cancer incidence was obtained through population cancer registries, or via a combination of three methods that included: health insurance records, cancer and pathology registries, and active follow-up (France, Germany, Naples, and Greece). Subjects were followed until cancer diagnosis (except non-melanoma skin cancer), emigration, death, or until the end of follow-up (dates varied between centers, from December 2004 to June 2010), whichever occurred first [127]. EC and EOC diagnoses were codded according to the International Classification of Diseases (ICD) as C54 and C56; respectively. The ICD is the standard diagnostic instrument for epidemiology, health administration and clinical purposes, and it is commonly used to record the incidence and prevalence of diseases and other health conditions [128].

4.3 Lifestyle information assessment

Information on lifestyle factors (smoking history, physical activity, education level, alcohol drinking), illnesses (i.e., diabetes), and menstrual and reproductive factors (i.e., OCs and HRT use, age at first menstrual period) were also collected at baseline using multiple-choice (close-ended) questionnaires and clinical examinations. Baseline anthropometry information was collected by interviewers in all centers and included: height, weight (self-reported in France), waist circumference and hip circumference. Umeå (Sweden) and France did not collect data on waist and hip circumference [118].

4.4 Dietary intake assessment

Food consumption in the EPIC cohort was measured at baseline by country-specific food intake questionnaires, which were designed to capture local dietary habits and to provide high compliance. In France, The Netherlands, Germany, Greece and Italy (except Naples) quantitative dietary questionnaires were used, but centers from United Kingdom and Naples used semi-quantitative food frequency questionnaires. Before the start of the study, these questionnaires were tested against biomarkers [119]. The use of different dietary assessment methods across the twenty three EPIC centers required that the dietary assessments obtained in each of the sub-cohorts could be comparable on an absolute scale. In order to calibrate the dietary questionnaires (DQs), additional dietary intake data was collected from a subset of the cohort by using a computer-assisted 24hDR (a standardized software in the language of each country named EPIC-SOFT). In total, 36,900 24hDR (~8% of the cohort) were collected and used as the reference calibration measurement [120]. Energy and nutrient intakes for the entire cohort were estimated using the standardized EPIC Nutrient Database (ENDB) [121].

4.5 Acrylamide exposure assessment

4.5.1 Dietary acrylamide intake assessment

In order to construct the EPIC-acrylamide database, the main source of information on acrylamide levels in foods was obtained from the European Commission Joint Research Centre's Institute for Reference Materials and Measurements (JRC-IRMM) that began collecting data from 2003 to 2006 [129]. Generally the IRMM database included chemical analysis of acrylamide levels in a wide variety of foods from Europe. The methods were mainly

based on either liquid chromatography (LC) or gas chromatography (GC) coupled to mass spectrometry (MS). The IRMM database includes acrylamide levels in foods mainly from Germany (77.6%), followed by Austria, Belgium, Finland, Greece, Ireland, Italy, The Netherlands, Spain, and the United Kingdom (all accounting for 10.3%), and from the food industry (12.1%) [130, 131]. Some food information that was not available at the JRC-IRMM database was taken from other country-specific databases (Swedish, Dutch, and Norwegian national databases), and also from the US FDA database (e.g., peanut butter) in order to have the best quality information.

As mentioned before, the EPIC study used different food intake assessments; thus different algorithms (but with the same ultimate aim) were applied to the raw intake data to calculate 'individual' intakes. 'Individual' intake was defined as the individual's average portion by frequency of consumption, rescaled for the same time unit (i.e., individual intake per day of a given food or food group). The information on acrylamide levels in food (obtained from the IRMM database and the supplemental databases) was merged and classified according to all food items from the ENDB [40, 47]. In those cases were an exact match was not possible, food consumption data from EPIC questionnaires were matched to similar items or to the mean of all items of the food group in the acrylamide IRMM and supplemental databases. Finally, all items were summed-up to obtain the total average daily acrylamide intake (µg/day) per individual (**Figure 11**).
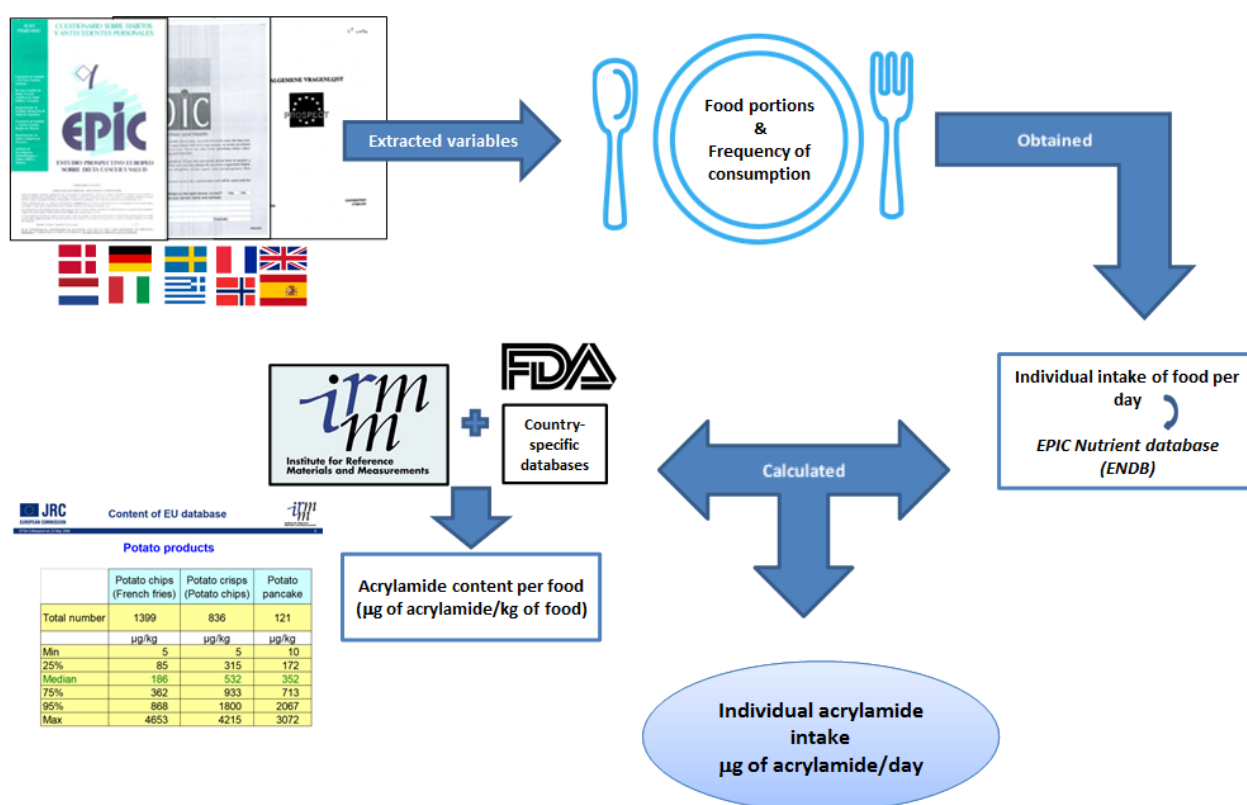

Figure 11. EPIC acrylamide database

4.5.2   Laboratory assays of hemoglobin adducts of acrylamide and glycidamide

Blood samples (serum, plasma, white blood cells, and red blood cells) were collected at baseline for 80% of the EPIC cohort. Red blood cells drawn from liquid nitrogen storage were

shipped from the EPIC biorepositories at IARC (Lyon), Denmark, and Sweden to the Centers for Disease Control and Prevention (CDC) Protein Biomarker Laboratory (Atlanta, US) to measure HbAA and HbGA. The methods used to analyze HbAA and HbGA have been published, and quality control/quality assurance activities in the laboratory are performed in compliance with the Clinical Laboratory Improvement Amendment (CLIA) [70]. Hemoglobin adducts were determined (in a randomized and blinded manner) by high performance liquid chromatography coupled with tandem mass spectrometry (HPLC/MS/MS) using ≈300 µL of hemolysed red blood cells per participant. Two independent adduct measures for each sample were performed. Levels of HbAA and HbGA were reported as the mean of these two measures relative to the amount of hemoglobin (pmol per g of Hb). The detection limits for HbAA and HbGA adducts are 3 and 4 pmol/g of Hb, respectively.

4.6 Statistical analyses

In EPIC, all statistical analyses are stratified by center or country (depending on the total number of cases), and age at recruitment (1-year categories). The reasons for these strata are: (a) the variation in dietary and cancer patterns in the EPIC data means it is important to include the variable center/country in the model, and (b) age is used as a primary time variable allowing left truncation of the data (then the risk of developing EC or EOC risk is only included in the model during the follow-up). The advantages of the strata inclusion are fewer parameters to estimate in the model, and the avoidance of violation of the proportional hazard (PH) assumption (in Cox models) for center/country and age at recruitment effects. A brief summary of the material and methods used in the articles can be found at **Table 9**.

4.6.1 Prospective cohort studies

Multivariate Cox Proportional hazard models were used to estimate hazard ratios (HR) and 95% confidence intervals (95% CI) for questionnaire-based acrylamide intake and EC or EOC risk.

Acrylamide consumption is correlated with total energy. For both studies, dietary acrylamide exposure was energy-adjusted using the residual method. This allowed to have a variable (acrylamide) that was independent of total energy intake. A simple linear regression model was run with acrylamide intake as the dependent variable and total energy intake as the independent variable to obtain the residual value (part of acrylamide intake that is not correlated with total energy intake). The energy-adjusted acrylamide intake for each EPIC participant was comprised of the sum of the residual (a; obtained from the regression model) and the expected acrylamide intake for a person with mean energy intake (x) (energy-adjusted acrylamide intake = a + x). The 'acrylamide intake' variable and the 'energy-adjusted acrylamide intake' variable share the same mean value, but have different deviation values (the 'energy-adjusted acrylamide intake' variable has less deviation than the 'acrylamide intake' variable) [132].

The continuous variable (average daily intake in 10 µg increments; 10 µg/day), and the categorical variable (quintiles of intake based on the distribution in the full EPIC cohort; µg/day) were evaluated in Cox models.

All models were stratified by country and age at recruitment (1-year categories), and adjusted for confounding variables and known risk factors for EC or EOC. Stratified analyses were performed with the aim to evaluate and control confounding, and to describe effect-measure modification by known risk factors or by factors that may affect the activity of the Cyp2e1. In order to avoid systematic errors (possible influences of pre-clinical disease on dietary habits) sensitivity analyses excluding all EC or EOC cases diagnosed during the first two years of follow-up were performed. In addition, analyses restricted to EC or EOC subtypes were also performed.

To evaluate dose-response trends the median value for each acrylamide quintile was estimated and included in a score test. Effect-measure modification was based on the continuous acrylamide intake variable, and was evaluated using a likelihood ratio test (LRT). All statistical tests were evaluated at the α-level 0.05. The Cox model assumes that the hazards are proportional to follow-up, and to test the PH assumption, Schoenfeld residuals were used [133]. Analyses were performed using SAS v.9.1 and STATA was used to test the PH assumption.

4.6.2 Cross-sectional study

Simple and multiple linear regression analyses were used to identify dietary and lifestyle determinants of HbAA, HbGA, sum of total adducts (HbAA+HbGA), and HbGA/HbAA ratio in 801 control non-smoking postmenopausal women from the two nested-case control studies of HbAA and HbGA levels and EC and EOC risk in EPIC.

All dependent and independent continuous variables were log-transformed ($\log_2$) to reduce skewness. First a correlation matrix was calculated to identify correlation between dietary and food group variables. Variables that were not correlated, matching factors, and lifestyle variables were candidates to be included in final models. Stepwise selection was used to build the four final models (one for each dependent variable).

Intraclass correlation coefficients (ICC) were estimated to evaluate the reproducibility of acrylamide adducts measurements. A Wilcoxon signed-rank test was used to assess differences in biomarker levels between subgroups (i.e., alcohol consumption). *R* square (R^2) values were used to describe the percent variation in biomarker levels explained by the independent variables. Partial R^2 values and Pearson's correlation coefficient, and Student's *t* statistics (and the corresponding *P*-value) were also estimated. All statistical tests were evaluated at the α-level 0.05. Analyses were performed using SAS v.9.1 and graphics were created using R v.3.1.

4.6.3 Nested-case control studies

To estimate odds ratios (OR) and 95% CI for both nested case-control studies (one control per case) on HbAA and HbGA levels and the risk of EC or EOC, unconditional logistic regression models were used, since the individual matching was broken in both studies. To avoid confounding from smoking and hormonal fluctuations, only those women who reported never smoking or who quit smoking ≥5 years before recruitment, and who were postmenopausal at blood draw were included. Conditional matched logistic regression was used as a sensitivity analysis for the association between biomarkers levels and EOC risk.

To improve normality of the distributions, HbAA, HbGA, HbAA+HbGA were logarithmically transformed. The HbGA/HbAA ratio followed a normal distribution. All four continuous variables were categorized into quintiles or quartiles (depending on the total number of cases) based on the exposure distribution in controls. All analyses were adjusted for matching variables, and confounding and risk factor variables. Analyses were also performed restricted to EC and EOC subtypes.

The reproducibility of acrylamide adducts measurements was assessed using ICC. Effect-measure modification was evaluated for established risk factors and for factors that may affect the activity of Cyp2e1 using the LRT test. To evaluate dose-response trends, the median for each biomarker variable was estimated and included in a score test. All statistical tests were evaluated at the α-level 0.05. Analyses were performed using SAS v.9.1.

4.7 Ethical considerations

All participants that accepted to be involved in the EPIC study signed an informed consent form, meaning that all information derived from diet and lifestyle questionnaires, as well as from blood sample extractions, could be used confidentially and anonymously for the study.

Furthermore, this study was authorized by the EPIC Steering Committee, the IARC Ethics Committee (IEC), and the 'Comité Ético de Investigación Clínica (CEIC) – Hospital Universitari de Bellvitge'. The IEC and the CEIC approvals are enclosed at the end of the document (**ANNEX; 9.2 Ethical Approvals**).

4.8 Funding

This thesis would not have been possible without the grants listed below. None of the funding agencies had a role in the study design, data acquisition, analysis and interpretation of data, or reporting results.

- Wereld Kanker Onderzoek Fonds (WCRF NL) [grant WCRF 2011/442]
- Health Research Fund (FIS) of the Spanish Ministry of Health [Exp PI11/01473]
- Red Temática de Investigación Cooperativa en Cáncer (RTICC) [RD12/0036/0018; RD06/0020/0091]

Table 9. Summary of the materials and methods used in the articles

	Publication 1	Publication 2	Publication 3	Publication 4	Publication 5
Design	Prospective cohort study		Cross-sectional study	Nested case-control study	
Subjects	Pre-, peri-, and postmenopausal women			Non-smoking and postmenopausal women	
Total number of subjects	301,113	325,006	801	768	751
Outcome	EC	EOC	Dietary & lifestyle determinants of HbAA and HbGA	EC	EOC
Number of cases	1,382 overall EC 627 type-I EC	1,191 overall EOC 582 serous invasive EOC	801 controls	383 overall EC & 385 controls 171 type-I EC &385 controls	334 overall EOC & 417 controls 191 serous invasive EOC & 417 controls
Exposure	Energy-adjusted acrylamide intake using the residual method		HbAA, HbGA, sum of total adducts (HbAA+HbGA), and HbGA/HbAA ratio		
Exposure variables	Continuous: average daily intake in 10 μg increments (10μg/day) Categorical: Quintiles of intake based on the distribution of acrylamide intake in the EPIC subcohort of women at baseline (μg/day)		*Dependent variables:* Log$_2$HbAA Log$_2$HbGA Log$_2$(HbAA+HbGA) HbGA/HbAA	Continuous: HbAA, HGA, HbAA+HbGA, HbGA/HbAA Continuous log-transformed: log$_2$HbAA, log$_2$HbGA, log$_2$(HbAA+HbGA) Categorical: continuous variables were categorized into quintiles based on the distribution in the control group	Categorical: continuous variables were categorized into quintiles based on the distribution in the control group. Biomarker quartiles were evaluated in stratified analysis.
Statistical Model	Simple and multivariate Cox proportional hazard models		Simple and multiple linear regression models	Unconditional and conditional logistic regression models	Unconditional and conditional logistic regression models
Statistical software	Analyses were performed using SAS v.9.1 and STATA was used to test the PH assumption		Analyses were performed using SAS v.9.1 and graphics were created using R v.3.1	Analyses were performed using SAS v.9.1	

EC, endometrial cancer; EOC, epithelial ovarian cancer; HbAA, hemoglobin adducts of acrylamide; HbGA, hemoglobin adducts of glycidamide; PH, Proportional Hazards