Online Data Supplements

Table of contents:

1 Online methods

Participants and covariates

Participants included in our study were women of Caucasian ancestry, defined as those who reported that they were "white Irish", "white British", or "other white" during all instances (N=257,210). The pattern of oral contraceptive use has differed dramatically between cultures and mixing participant with different backgrounds could have led to misleading results. Unfortunately, the non-Causation groups had too few individuals to be analyzed separately. Individuals who either answered "do not know" or preferred not to answer the question: "Have you ever taken the oral contraceptive pill" were removed from the analyses (N=533) as well as women who have requested to be removed from the UK Biobank (UKB). After this filtering, 256,661 women remained for our study. Information on oral contraceptive use was taken from the initial assessment visit. This information originates from the touchscreen questionnaire, data fields: 2794 (age started oral contraceptive pill), 2804 (age when last used oral contraceptive pill), and 2784 (ever taken oral contraceptive pill). Duration of use was estimated as the difference between age when oral contraceptives were initiated and age of last use. For participants who were still using oral contraceptives (N=6,698), the age when last used was set to the age when they visited the assessment center.

Assessment of disease status and covariate information

Our study included both self-reported and hospital registered data (Supplementary Table S2). A larger proportion of self-reported data was collected at first assessment (2006-2010) and the registry data is less complete before 1995. For the period between 1995 and 2006, the overlap between self-reported and registry data was large, and the self-reported and registry-derived data of first diagnosis agreed well (Supplementary Figure S2). We can therefore consider the self-reported data used to address questions retrospectively to be accurate to address questions prospectively.

Information on year of birth, sex and Townsend deprivation index (TDI - used as a proxy for socioeconomic status), were obtained from the local National Health Service Primary Care Trust register immediately prior to the participant joining UKB. Body mass index (BMI) was constructed from height and weight measures, which were obtained at the initial assessment. Smoking status was gathered from the touchscreen questionnaire and participants were classified as never, previous, occasionally or current smokers. Menopausal status was obtained from the initial assessment. Women were asked through the touchscreen questionnaire: "have you had your menopause (periods stopped)?". The participant could answer, "yes", "no", "not sure – had a hysterectomy", "not sure – other reason" and "prefer not to answer". Participants who had answered "prefer not to answer" were excluded from all analyses that included menopausal status. In the cumulative-risk analyses, menopausal status was incorporated as a factor covariate with four levels (see above). All women, except those that had already answered "not sure – had hysterectomy", on menopause status, were further asked about whether they had undergone hysterectomy, and a second variable indicating if the participants had answered that they had hysterectomy at any of the two questions were also included as a covariate. Number of live births, age at menarche (age when periods started) and hormonal replacement treatment (HRT) was obtained from the self-reported questionnaire data from the initial assessment. Participants answering "do not know", "prefer not to answer" or "none of the above" were set as missing and excluded in all analyses of the respective covariate. All quantitative variables were analyzed as quantitative, and age were expressed in whole years.

Statistical analyses

Logistic regression

Logistic regression models were used to calculate odds ratio (OR) and 95% confidence interval (CI) between ever and never users, for each cancer diagnosis. All cancers were also analyzed by stratifying for duration of oral contraceptive use into six different intervals (<2, ≥2 to <5, ≥5 to <10, ≥10 to <15, ≥15 to <20, ≥20 years) and compared to never users, using the full model (model 4). Participants who had not reported when they started with oral contraceptives were excluded from the duration of use analyses. For participants who were still using oral contraceptives, the duration time was estimated to be the time between starting of oral contraceptive use and the initial visit to the assessment center.

Several sensitivity analyses were also performed:

A. To evaluate the influence of the covariates, the effect of each specific covariate on cancer diagnosis was tested using a simple logistic model, i.e. without taking oral contraceptive use into account. The covariates included smoking status, age, BMI, TDI, year of birth, age at menarche, menopausal and hysterectomy status, number of live births and HRT. Most of the tested covariates were associated with a cancer diagnosis (Supplementary Table 2). We therefore also included each covariate, one-by-one, when estimating the effect of oral contraceptive use on each cancer (Table 2, main article).
B. We further analyzed four models, consisting of different sets of covariates, when estimating the effect of oral contraceptive use on cancer risk, to explore potential confounding (Table 2, main article). The full model (model 4), which was used for generating the main results in this article, included age, BMI, TDI, year of birth, smoking status, age at menarche, HRT, number of live births, as well as menopausal and hysterectomy status.
C. Since the cancer incidence rates increase dramatically by age (Figure 1, main article), the large number of cases that are not related to oral contraceptive use, might dilute the OR when following participants until a high age, similar to what has been suggested for relative risks previously(1). We therefore performed additional logistic regression analyses, in which only cancer diagnoses prior to a certain year of age were considered (Table 3, main article). For example, with a follow-up to 55 years of age, participants who were diagnosed at age 55 or earlier were set to cases, while participants who did not have a diagnosis at age of 55 were set to controls. This was done for the set of follow up-ages: 35, 40, 45, 50, 55, 60, 65, 70, 75, and 80 years.
D. To further discriminate between cancer prevalence at recruitment of the UKB cohort, and new cancer diagnoses after recruitment, we re-estimated the ORs between never and ever users, as well as between never and previous users, for all incident cases after initial visit to the assessment center. Here, we excluded all participants that reported that they have been diagnosed with the cancer under investigation before the date of the initial visit to the assessment center, while incident cases were identified, through the register data, as cancer diagnoses with first date of diagnosis after the initial visit to the assessment center. By adopting this procedure, we reduce the risk for reversed causation, in the sense that a prior cancer diagnosis might influence the choice of using oral contraceptives. The data were analyzed using logistic regression modelling, with the same covariates as those included in the full model (model 4; see Table 2, main article).

Cox regression

Time-varying covariates and time-dependent coefficients
Both time-dependent coefficients, as realized by stratification of follow-up time, and time-varying covariates were incorporated within the Cox model, by the counting process formulation (2). Some covariates were fixed and equal in both during and after use analyses, such as age at menarche and year-of-birth. BMI and TDI were also unchanged between the analyses and given at time of assessment. Age was adjusted for by including age at entry in the models, which implied age at first use of oral contraceptives for the during-use analysis and age at last use of oral contraceptives for the after-use analyses. Covariates with information on changes within the follow-up time (smoking, menopause, hysterectomy, and HRT) were modelled as time-varying. In the logistic regression, smoking status was modelled as a factor with four levels: never, former, present, and occasional. In the Cox regression, smoking status was not fixed, but could change with time as given by the ages when started (data-fields 3436 and 2867) and stopped smoking (data-field: 2897). Occasional smokers were assumed not to have changed their smoking status until time of assessment. Menopause was also modelled as a factor with four levels in the logistic regression. In the Cox regression, the age at which change of menopausal status occurred was extracted, either from data-field 3581 (age at menopause), 2824 (age at hysterectomy), for women who were unsure about menopausal status due to hysterectomy, or from data-field 21003 (age at assessment), for individuals who were unsure about menopausal status due to other (unknown) reasons. The age distribution (at assessment) for individuals who were unsure about menopausal status due to other (unknown) reasons closely resembled the age distribution at menopause for individuals who stated that they have gone through menopause (Supplementary Figure S4), wherefore we considered these women to enter menopause at age of assessment. Hysterectomy (womb removed, data-field 3591) was modelled as a dichotomous, time-varying variable (no/yes), with age at hysterectomy taken from data-field 2824. In the Cox regression, ages for start and stop with HRT was taken from data-fields 3536 (age started HRT) and 3546 (age last used HRT). Individuals who stated that they still take HRT in data-field 3546 were assumed to do so until time of assessment. Individuals who did not know or preferred not to answer in any of the above covariates were excluded from the analyses.

The time dependent covariates were incorporated in the Cox model using the tmerge function in the R package survival (v 2.44.1.1). The effect of oral contraceptive use was also allowed to vary with time. This was accomplished by describing the data in counting-process format (2), where the data were split in an additional number of cut times, using the survSplit function (v 2.44.1.1) in R. The default cut times for the during-use analysis were set to 2, 5, 10, 15, and 20 years (same as the duration of use analyses in the logistic regression analysis), while the cut times for after-use analysis were set to 2, 5, 10, 15, 20, 25, 30, and 35 years.

Start and stop ages for never users and estimation of standard errors
All Cox regression models were fitted using the coxph function in the survival package (v 2.44.1.1) in R. A robust estimate of the standard error was computed by adding a cluster(id) term in the models. Since never users did not have start and stop ages for oral contraceptive use, each never user was assigned a pair of start and stop age, randomly selected from an ever user with matched age at assessment, when entering UK Biobank. This random selection was performed with replacement. To reduce the additional variance from this procedure, we computed the mean estimate of log (HR) from 100 runs, in which each never user was

(randomly) assigned different pairs of start and stop ages. The total standard error of the mean estimate is then given by

$$se_{\text{tot}} = \sqrt{\frac{\sum_{i=1}^{n} se_i^2 + \frac{1}{n-1}\sum_{i=1}^{n}\left(\beta_i - \bar{\beta}\right)^2}{n}},$$

where β_i and se_i denote the estimates of log(HR) and the corresponding robust standard error for run i, respectively, $\bar{\beta}$ denotes the mean estimate of log(HR), and $n = 100$ is the number of runs. As a reference, the between-variance (second sum) contributed less than 0.6% of the total variance in all analyses with $n = 100$.

Age-matched tests and sensitivity analyses for Cox regression

Several sensitivity analyses were performed, to assess the robustness of the main results. Firstly, individuals were only followed until age at hysterectomy or age at bilateral oophorectomy, whichever came first, if they had hysterectomy or bilateral oophorectomy before the end of normal follow-up. Secondly, we considered the potential effect modification of giving birth, where individuals during oral contraceptive use were followed, either until they stopped using oral contraceptives, or until their first live birth, whichever came first.

It is important to consider the differences between study designs and the potential impact it can have on the results. A clear limitation in many studies is lack of an age-matched test statistic. Previous studies have often used Poisson regression (3) in age bands of, for example, 5 years. It is well known that cancer incidence rates increase with age in a non-linear manner (1), which is also true in our study cohort (see Figure 1, in the main article) and a small difference in the age-distribution between never/current/previous oral contraceptive users can potentially have a dramatic effect on the results. We therefore performed additional analyses to further test the possible issue of an age difference between ever and never users. In a third sensitivity analysis, the distribution of ages for ever- and never users were matched and made equal by re-sampling, in addition to the age-matching where never-users were assigned start and stop ages for oral contraceptive use, to further eliminate the possibility of a residual effect modification due to a difference in ages. Fourthly, to assess whether a linear adjustment for age was adequate in our main analyses, we used age, instead of follow-up time, as primary time scale (4). The results from these two sensitivity analyses were very similar to those of the main analyses. This suggests that the age-matching and linear adjustment for age give adequate results, which are not biased by any residual difference in age distribution between cases and controls. To our knowledge, our study is the first to consider age-matched groups when investigating the effect of oral contraceptive use on ovarian, endometrial and breast cancer. Lastly, in ever users, there was a correlation between age at menopause and age at stopping of oral contraceptives which was not present in never users, since never users were assigned start and stop times independently of their menopausal status. To assess the effect of this difference between ever and never users, the difference in age between menopause and stopping of oral contraceptives was matched by re-sampling, similar to above, to being equal in distribution between ever and never users, before fitting the Cox regression models.

References to supplementary files

1. Breast cancer and hormonal contraceptives: collaborative reanalysis of individual data on 53 297 women with breast cancer and 100 239 women without breast cancer from 54 epidemiological studies. Lancet. 1996;**347**:1713–27.
2. Andersen PK, Gill RD. Cox's Regression Model for Counting Processes: A Large Sample Study. Ann Stat. 1982;
3. Mørch LS, Skovlund CW, Hannaford PC, Iversen L, Fielding S, Lidegaard Ø. Contemporary Hormonal Contraception and the Risk of Breast Cancer. N Engl J Med. 2017;**377**:2228–39.
4. Cologne J, Hsu W, Abbott RD, Ohishi W, Grant EJ, Fujiwara S, et al. Proportional Hazards Regression in Epidemiologic Follow-up Studies An Intuitive Consideration of Primary Time Scale. Epidemiology. 2012;**23**:565–73.
5. Calle EE, Heath CW, Miracle-McMahill HL, Coates RJ, Liff JM, Franceschi S, et al. Breast cancer and hormonal contraceptives: Collaborative reanalysis of individual data on 53 297 women with breast cancer and 100 239 women without breast cancer from 54 epidemiological studies [Internet]. Lancet. 1996. page 1713–27.
6. Marchbanks PA, McDonald JA, Wilson HG, Folger SG, Mandel MG, Daling JR, et al. Oral Contraceptives and the Risk of Breast Cancer. N Engl J Med. Massachusetts Medical Society ; 2002;**346**:2025–32.
7. Iversen L, Fielding S, Lidegaard Ø, Mørch 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. British Medical Journal Publishing Group; 2018;**362**:k3609.
8. Kumle M, Weiderpass E, Braaten T, Persson I, Adami H-O, Lund E. Use of oral contraceptives and breast cancer risk: The Norwegian-Swedish Women's Lifestyle and Health Cohort Study. Cancer Epidemiol Biomarkers Prev. Cancer Epidemiol Biomarkers Prev; 2002;**11**:1375–81.
9. Rosenblatt KA, Gao DL, Ray RM, Nelson ZC, Wernli KJ, Li W, et al. Oral contraceptives and the risk of all cancers combined and site-specific cancers in Shanghai. Cancer Causes Control. NIH Public Access; 2009;**20**:27–34.

2 Supplementary Tables

Supplementary Table S1. A selection of previous studies aiming to estimate the effect of oral contraceptive use on risk of breast cancer*

Reference	Cohort/Study design	Year of birth, age at recruitment, Study years	Year at diagnosis or end of follow up	Relative Risk [95% confidence interval]
Collaborative Group on Hormonal Factors in Breast Cancer, 1996** (5)	53,297 breast cancer cases and 100,239 controls	The median age at diagnosis of breast cancer was 49 years. 33% were 55 years and older	Most cancers were diagnosed during the 1980s, median year 1984	Current users RR= 1.24 [1.15-1.33]
Marchbanks et al. 2002 (6)	4,575 breast cancer cases and 4,682 controls	35 to 64 years old at diagnosis	Diagnosed between 1994 and 1998	Current users RR = 1.0 [0.8 - 1.3]
Hannaford et al. 2007 (7)	1,339 breast cancer cases and 44,661 controls***	Mean age at recruitment (1968-1969): 29 (SD: 6.6)	Year: 2004. Median age of about 65 years	Ever users: RR = 0.98 [0.87 - 1.10]
Mørch et al. 2017 (3)	Prospective cohort: 1.8 million women, 11,517 cases of breast cancer	Between 15 and 49 years in 1995	Year 2012. Age < 50	Current/recent users RR = 1.20 [1.14 - 1.26]
Kumle et al. 2002 (8)	Prospective cohort: 103,027 women 1,008 primary invasive breast cancers	Aged 30-49 years at recruitment (1991/1992)	Year 1999. Median age about 39-58 years in 19999	Current/recent users RR= 1.6 [1.2-2.1]
Rosenblatt et al. 2008 (9)	Prospective cohort: 267,000 women and 1,749 breast cancers cases	Born between 1925 and 1958. Study started in 1989–1991	Year 2000. 42-75 years of age at end of follow up	Ever users RR= 0.90 [0.78–1.03)]

* Only studies with at least 1000 breast cancer cases are included
** Meta-analysis of 54 studies conducted over 25 years
*** 23,000 current oral contraceptive users, and 23,000 never users were included

Supplementary Table S2. Identification of cancer cases in the UK biobank database.

Cancer	Type of data	Data-fields*	N** per data-field	N total*** (N unique****)	Coding	Explanation of coding
Breast	Registry data (ICD-10)	40006/41202/41204/40001/40002	14116/13143/3606/897/69	15704	C50	Malignant neoplasm of breast
	Registry data (ICD-9)	40013/41203/41205	1728/193/11	1771 (761)	174	Malignant neoplasm of female breast
	Self-reported - verbal interview	20001	11334	11334 (1274)	1002	Breast cancer
Ovarian	Registry data (ICD-10)	40006/41202/41204/40001/40002	1235/1169/645/394/15	1606	C56	Malignant neoplasm of ovary
	Registry data (ICD-9)	40013/41203/41205	147/11/0	148 (86)	183	Malignant neoplasm of ovary and other uterine adnexa
	Self-reported - verbal interview	20001	836	836 (274)	1039	Ovarian cancer
Endometrial	Registry data (ICD-10)	40006/41202/41204/40001/40002	1718/1589/281/111/8	1961	C541	Endometrium
	Self-reported - verbal interview	20001	1251	1251 (501)	1040	Uterine/endometrial cancer

* Explanations to data fields:

In-patient hospital admissions, Summary Information (diagnoses)
41202 Diagnoses - main ICD10
41204 Diagnoses - secondary ICD10
41203 Diagnoses - main ICD9
41205 Diagnoses - secondary ICD9
Death register - Death records contains both primary cause of death and contributory cause of death, provided to UK Biobank from national death registries
40001 Underlying (primary) cause of death ICD10
40002 Contributory (secondary) causes of death ICD10
Cancer register - The earliest date for which cancer registry data is available is 20 September 1957, therefore cancer diagnoses prior to this date are not included here
40013 Type of cancer ICD9
40006 Type of cancer ICD10
Cancer code, self-reported - verbal interview by a trained nurse
20001 Cancer type

** N is the number of cancer incidences identified within each data field
*** Total N is the number of cancers identified in any of the data fields in the data-field column
**** N unique is the number of cases that was not identified in the definitions on the line(s) above for the same cancer (i.e. ICD-9 not found in ICD-10 or self-reported not found in registry data)

Supplementary Table S3. Identification of covariates in the UK biobank database and their Odds ratio [OR] and 95% confidence interval [CI]) on odds of cancer.

Covariates	Data-fields	Registry	N**	Breast cancer OR (95% CI), P-value	Ovarian cancer OR (95% CI), P-value	Endometrial cancer OR (95% CI), P-value
Age	21003	Registry	257194	1.05 (1.05-1.05), <0.001	1.05 (1.04-1.05), <0.001	1.08 (1.07-1.08), <0.001
Year of birth	34	Registry	257194	0.95 (0.95-0.96), <0.001	0.95 (0.95-0.96), <0.001	0.93 (0.92-0.93), <0.001
BMI (body mass index)	21001	Measured	256152	1.01 (1.00-1.01), <0.001	1.02 (1.01-1.03), <0.001	1.08 (1.08-1.09), <0.001
TDI (Townsend deprivation index)	189	Derived variable	256894	0.99 (0.99-1.00), 0.0024	1.00 (0.99-1.02), 0.824	1.02 (1-1.03), 0.0117
Age at menarche (first period)	2714	Self-reported	249889	0.97 (0.96-0.98), <0.001	0.97 (0.94-1.00), 0.034	0.90 (0.88-0.92), <0.001
Smoking – Never	20116/1239	Self-reported	150127	Reference factor***	Reference factor***	Reference factor***
Smoking – Current	20116/1239	Self-reported	17767	0.92 (0.87-0.98), 0.015	1.10 (0.92-1.31), 0.308	0.83 (0.7-0.98), 0.0320
Smoking – Occasional	20116/1239	Self-reported	36742	1.12 (1.07-1.17), <0.001	1.07 (0.94-1.22), 0.281	0.83 (0.73-0.94), 0.00284
Smoking – Previous	20116/1239	Self-reported	51683	1.18 (1.14-1.23), <0.001	1.15 (1.03-1.29), 0.0132	1.07 (0.97-1.19), 0.153
Number of live births	2734	Self-reported	256993	0.97 (0.96-0.98), <0.001	0.89 (0.85-0.92), <0.001	0.95 (0.91-0.98), 0.00113
Hormone replacement treatment (HRT) – Ever	2814	Self-reported	100677	Reference factor***	Reference factor***	Reference factor***
Hormone replacement treatment (HRT) – Never	2814	Self-reported	155838	1.11 (1.07-1.14), <0.001	1.82 (1.66-1.99), <0.001	1.27 (1.17-1.37), <0.001
Had menopause - No	2724	Self-reported	58457	Reference factor***	Reference factor***	Reference factor***
Had menopause - Yes	2724	Self-reported	157728	2.80 (2.67-2.95), <0.001	3.02 (2.56-3.56), <0.001	3.98 (3.4-4.66), <0.001
Had menopause – Not sure – other*	2724	Self-reported	10859	2.67 (2.45-2.90), <0.001	1.80 (1.32-2.46), <0.001	1.81 (1.34-2.44), <0.001
Had menopause – Not sure – hysterectomy*	2724	Self-reported	29850	2.23 (2.09-2.38), <0.001	5.80 (4.84-6.95), <0.001	4.93 (4.12-5.89), <0.001
Hysterectomy - Ever	2724/3581	Self-reported	48537	Reference factor***	Reference factor***	Reference factor***
Hysterectomy - Never	2724/3581	Self-reported	208418	1.20 (1.16-1.24), <0.001	5.29 (4.84-5.79), <0.001	5.00 (4.62-5.41), <0.001

* Participants that were not sure about menopausal status were divided into those that were unsure because of hysterectomy and those unsure for other, unspecified, reasons
** N – number of individuals with covariate data available for quantitative variables, and number of individuals within each level for factors
*** Reference factor – is the factor level that the other factor(s) is compared to in the logistic regression

Supplementary Table S4. Fixed (time-independent) hazard ratios of oral contraceptives, during and after use.

Type of analysis	Breast cancer Events/Person-years*	Breast cancer HR (95% CI), P-value**	Ovarian cancer Events/Person-years*	Ovarian cancer HR (95% CI), P-value**	Endometrial cancer Events/Person-years*	Endometrial cancer HR (95% CI), P-value**
During use	488/2,367,389	0.87 (0.68-1.13), 0.30	68/2,368,186	0.46 (0.26-0.85), 1.2×10^{-2}	102/2,367,990	0.74 (0.45-1.23), 0.25
After use	8,975/5,334,099	1.00 (0.95-1.06), 0.92	630/5,406,519	**0.62 (0.52-0.75), 1.2×10^{-6}**	911/5,403,234	**0.61 (0.52-0.71), 8.5×10^{-11}**

In the Cox regression analysis, associations between oral contraceptive use and cancer are measured by the hazard ratio (HR). In the during-use analyses, the associations were estimated between first and last use of oral contraceptives (follow-up time), and compared users to non-users, while in the after-use analyses, associations were estimated between last use and time of assessment (visiting an assessment center).
* Median number of events (never and ever users) and corresponding median person-years, based on 100 iterations
** P-values in bold are significant after adjusting for multiple testing using Bonferroni method (3 cancers × 2 tests; P<0·0083)

Supplementary Table S5. Fixed (time-independent) hazard ratios of oral contraceptives, during and after use, with age as primary time scale in the Cox regression modelling.

Type of analysis	Breast cancer (events/person-years)	Breast cancer HR (95% CI), P-value**	Ovarian cancer Events/person-years*	Ovarian cancer HR (95% CI), P-value**	Endometrial cancer Events/person-years*	Endometrial cancer HR (95% CI), P-value**
During use	489/2,367,702	0.89 (0.69-1.16), 0.39	68/2,367,986	0.47 (0.25-0.86), 1.5×10^{-2}	101/2,368,218	0.77 (0.46-1.29), 0.33
After use	8,973/5,333,139	0.97 (0.92-1.02), 0.23	628/5,405,946	**0.63 (0.52-0.76), 2.3×10^{-6}**	914/5,402,209	**0.60 (0.52-0.70), 4.1×10^{-11}**

This analysis is analogous to the analysis presented in Supplementary Table S4, but with age as primary time scale. This is done to naturally adjust for the potential nonlinear effect of age, without having to include nonlinear age terms as covariates. These results are based on 25 iterations.
* Median number of events (never and ever users) and corresponding median person-years, based on 25 iterations
** P-values in bold are significant after adjusting for multiple testing using Bonferroni method (3 diseases × 2 tests; P<0·0083)

Supplementary Table S6. Time-dependent hazard ratios of oral contraceptives during use.

Duration of use	Breast cancer events*	Breast cancer HR (95% CI)	Ovarian cancer events*	Ovarian cancer HR (95% CI)	Endometrial cancer events*	Endometrial cancer HR (95% CI)
< 2 years	10	0.40 (0.10-1.52)	9	0.33 (0.07-1.48)	13	0.82 (0.23-3.00)
≥ 2 - < 5 years	13	0.51 (0.15-1.68)	4	0.25 (0.04-1.50)	14	0.63 (0.16-2.45)
≥ 5 - < 10 years	41	0.58 (0.28-1.20)	18	0.57 (0.17-1.93)	26	0.69 (0.26-1.83)
≥ 10 - < 15 years	89	1.29 (0.69-2.39)	15	0.66 (0.18-2.41)	21	1.30 (0.38-4.40)
≥ 15 - < 20 years	73	0.72 (0.41-1.26)	9	0.61 (0.12-3.20)	23	1.18 (0.29-4.75)
≥ 20 years	240	1.07 (0.71-1.61)	10	0.54 (0.10-2.78)	6	0.37 (0.08-1.67)

In the Cox regression analysis, associations between oral contraceptive use and cancer are measured by the hazard ratio (HR). The association is estimated between first and last use of oral contraceptives.
* Total number of events (never and ever users) in the last iteration

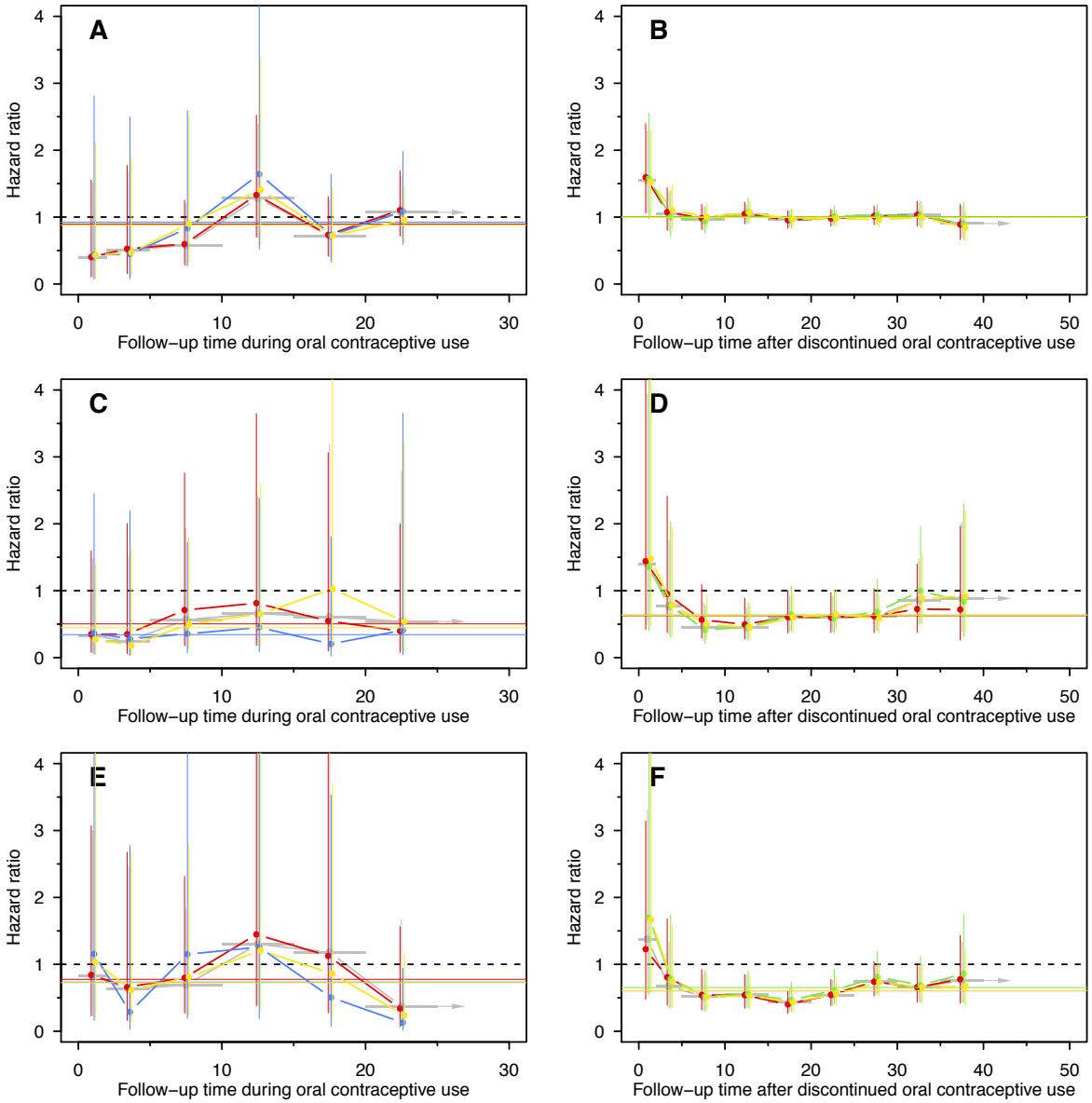
Supplementary Table S7. Time-dependent hazard ratios of oral contraceptives after last use.

Time after stopping	Breast cancer Events*	Breast cancer HR (95% CI)	Ovarian cancer Events*	Ovarian cancer HR (95% CI)	Endometrial cancer Events*	Endometrial cancer HR (95% CI)
< 2 years	267	1.55 (1.06-2.28)	37	1.39 (0.49-3.94)	47	1.38 (0.57-3.30)
≥ 2 - < 5 years	394	1.04 (0.79-1.37)	40	0.77 (0.34-1.75)	43	0.67 (0.34-1.34)
≥ 5 - < 10 years	857	0.96 (0.80-1.14)	54	0.46 (0.26-0.82)	87	0.52 (0.32-0.85)
≥ 10 - < 15 years	1,317	1.03 (0.90-1.19)	73	0.45 (0.27-0.75)	102	0.55 (0.36-0.84)
≥ 15 - < 20 years	1,606	0.96 (0.85-1.09)	110	0.60 (0.39-0.93)	121	0.43 (0.29-0.64)
≥ 20 - < 25 years	1,777	0.99 (0.88-1.11)	111	0.60 (0.39-0.91)	154	0.54 (0.39-0.76)
≥ 25 - < 30 years	1,483	1.02 (0.90-1.15)	89	0.62 (0.40-0.96)	178	0.75 (0.55-1.03)
≥ 30 - < 35 years	806	1.03 (0.88-1.21)	68	0.86 (0.50-1.48)	102	0.65 (0.44-0.97)
≥ 35 years	271	0.90 (0.70-1.16)	23	0.89 (0.39-2.01)	50	0.76 (0.43-1.33)

In the Cox regression, associations between oral contraceptive use and cancer are measured by the hazard ratio (HR). The HR is estimated between age at last use of oral contraceptives and age at assessment, when entering UKB.
* Total number of events (never and ever users) in the last iteration

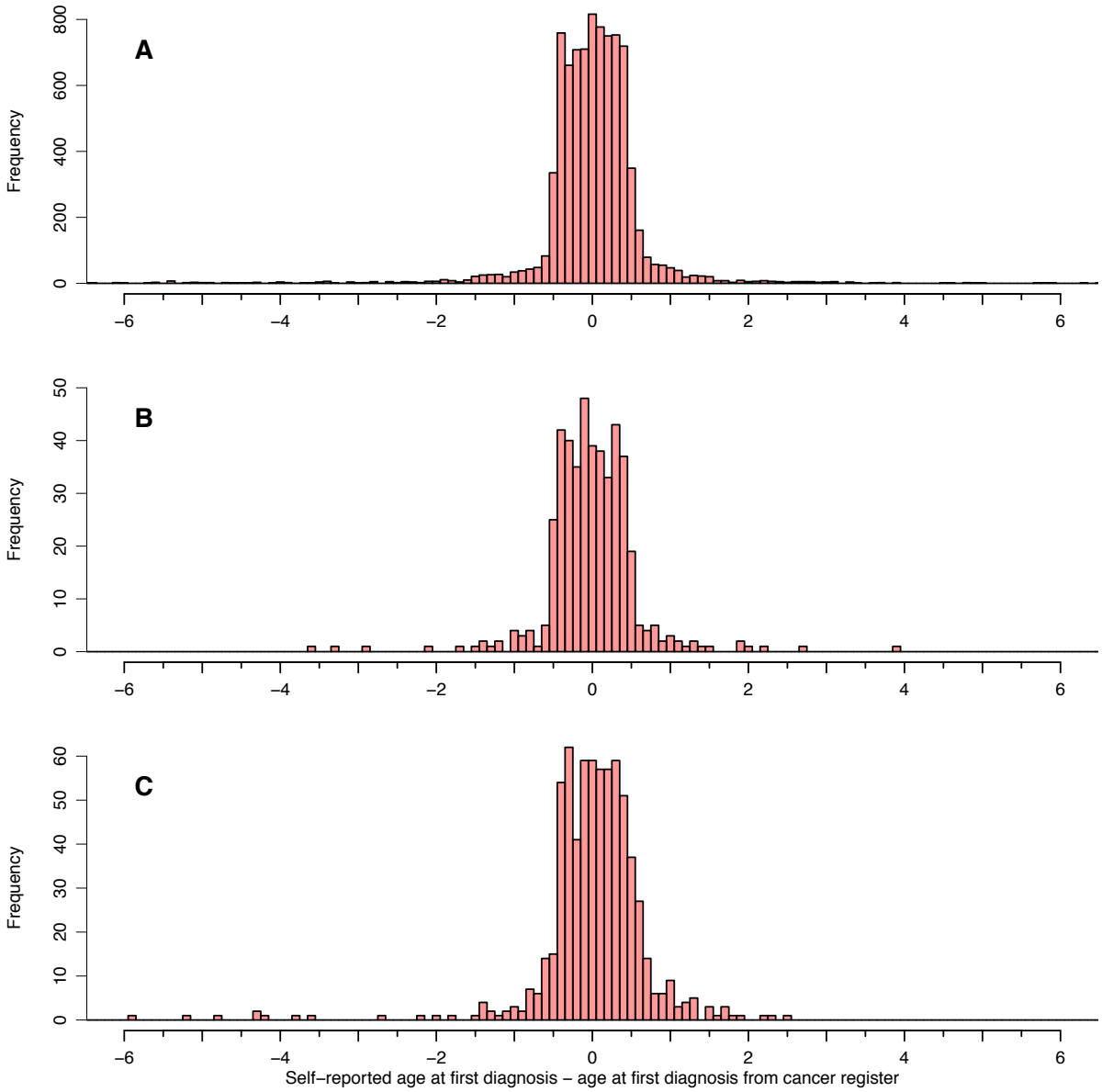
3 Supplementary Figures

Supplementary Figure S1. Sensitivity analyses for the associations between oral contraceptive use and cancer in the Cox regression



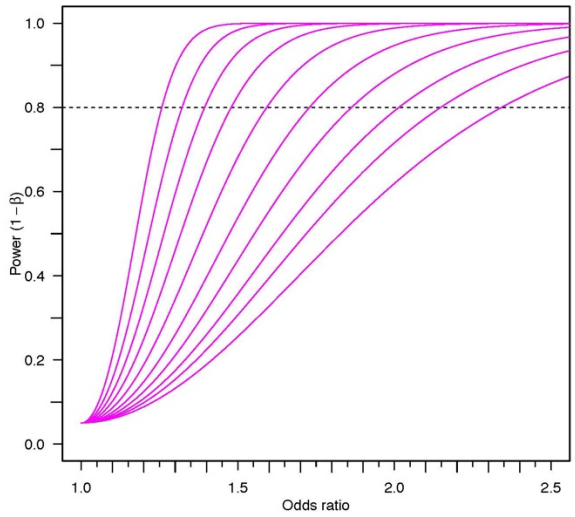
Cox regression during use (A, C, E) and after discontinued use (B, D, F) for breast cancer (A, B), ovarian cancer (C, D) and endometrial cancer (E, F). The points are the time-dependent HR with its 95% CI as vertical lines. Follow-up until hysterectomy or bilateral oophorectomy are denoted in red. Follow-up until first live birth are denoted in blue. Green denotes results from a re-sampling, where the excess fraction of users, relative to non-users, who stopped using oral contraceptives because of entering menopause is removed. Yellow denotes results from a re-sampling of the age distributions of ever- and never-users, to remove any residual differences in the ages between the two groups. Finally, grey denote the results from the main analysis (Figure 3, main article), where the thick horizontal grey bars represent the time strata that are contributing to each point. The solid, colored lines denote the fixed, time-independent HR corresponding to respective analysis.

Supplementary Figure S2. Discrepancy between self-reported data and cancer register



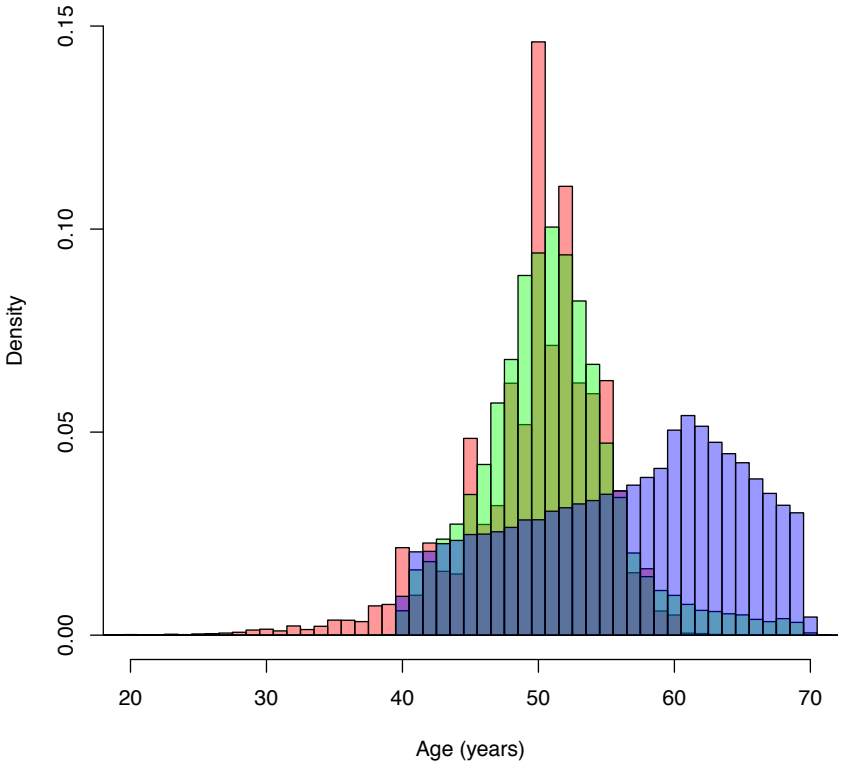
Histograms show the difference in age at first diagnosis between self-reported (interpolated) data (UKB) and cancer register data (ICD10) for all individuals where both ages are reported, for A) breast cancer, B) ovarian cancer, and C) endometrial cancer. The discrepancy is small, with an absolute median difference of less than 0.05 years and an interquartile range of the difference of 0.6 years for all three cancers.

Supplementary Figure S3. Power calculations



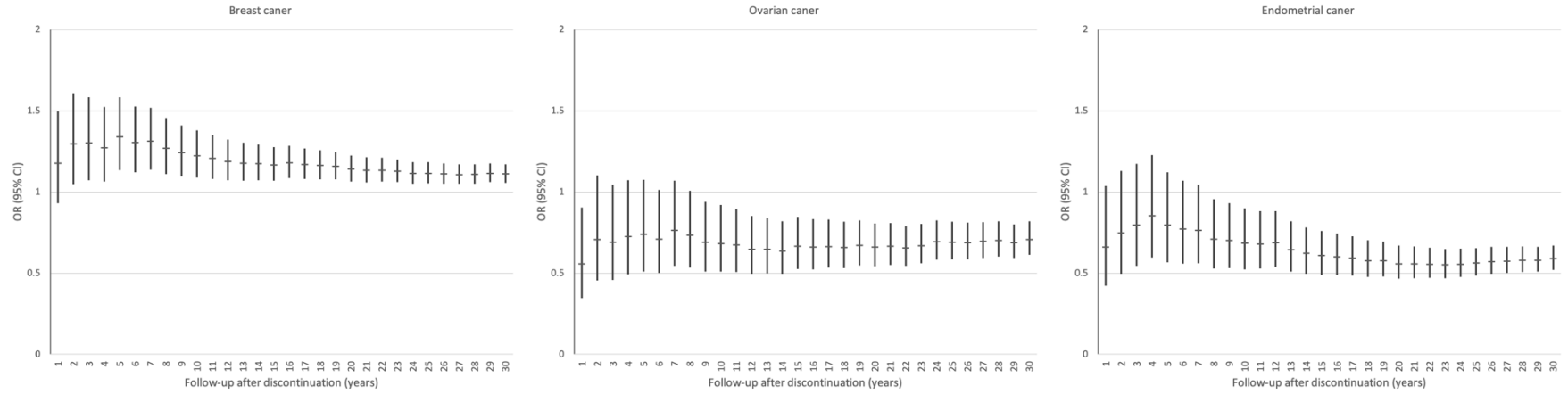
The curves show the expected power as a function of OR, given different number of disease cases. The numbers of controls are fixed to 50,000 never users and 200,000 ever users. The numbers of cases of never users are fixed for each curve to, from right to left: 12, 15, 18, 23, 30, 42, 60, 84, 120, 180, while the numbers of cases of ever users are allowed to vary. At OR = 1 (null association), the number of cases of ever users for each curve are a factor of four larger than the never-user cases.

Supplementary Figure S4. Distributions for age at menopause



The distribution in age at menopause (last menstrual period) for individuals who stated that they have had menopause are shown in red, while the green histogram indicates the distribution in age at assessment for individuals who stated that they were not sure about their menopausal status (other reason). The distribution in age at assessment for all women included in this study are shown in blue, for reference. Clearly, there is a strong resemblance between the red and the green histograms, wherefore we take the age at assessment for women who stated "not sure – other reason" to be their age at menopause.

Supplementary Figure S5. Sensitivity analyses for different follow-up years after oral contraceptive discontinuation.



ORs for current and previous users for different follow-up years after discontinuation. Each OR (and 95% confidence interval) is the odds for developing cancer during use or within the follow-up time (in years as indicated by the x-axis) after discontinuation, in relation to the odds among never users. The first point (1 year) represents the OR among individuals who had a reported cancer diagnosis up to 1 year after last oral contraceptive use. A total of 10,927 incidences of breast, 1048 incidences of ovarian and 1220 incidences of endometrial cancer were included in the analysis, after 30 years of follow-up.