Human Reproduction, Vol.34, No.5 pp. 894–902, 2019
Advanced Access publication on March 13, 2019 doi:10.1093/humrep/dez018

human reproduction

ORIGINAL ARTICLE *Reproductive epidemiology*

Risk of cancer in infertile women: analysis of US claims data

Gayathree Murugappan[1,*], Shufeng Li[2], Ruth B. Lathi[1], Valerie L. Baker[3], and Michael L. Eisenberg[1,2]

[1]Division of Reproductive Endocrinology and Infertility, Department of Obstetrics and Gynecology, Stanford Hospital and Clinics, Stanford, CA, USA [2]Department of Urology, Stanford University School of Medicine, Stanford, CA, USA [3]Division of Reproductive and Endocrinology, Department of Gynecology and Obstetrics, Johns Hopkins University School of Medicine, Baltimore, MD, USA

*Correspondence address. Stanford Department of Reproductive Endocrinology and Infertility, 1195 W Fremont Ave, Sunnyvale, CA 94087, USA. Tel: +1-650-498-7911; E-mail: gayathreem@gmail.com

Submitted on September 19, 2018; resubmitted on January 11, 2019; accepted on February 8, 2019

STUDY QUESTION: Is female infertility associated with higher risk of cancer?

SUMMARY ANSWER: Although absolute risks are low, infertility is associated with higher risk of cancer compared to a group of non-infertile women.

WHAT IS KNOWN ALREADY: Infertile women are at higher risk of hormone-sensitive cancers. Information on risk of non-gynecologic cancers is rare and conflicting, and the effect of pregnancy on these risk associations is known for only a minority of cancer types.

STUDY DESIGN, SIZE, DURATION: Retrospective cohort analysis between 2003 and 2016 using an insurance claims database.

PARTICIPANTS/MATERIALS, SETTING, METHODS: In all, 64 345 infertile women identified by infertility diagnosis, testing or treatment were compared to 3 128 345 non-infertile patients seeking routine gynecologic care. Women with prior diagnosis of cancer or within 6 months of index event were excluded. Main outcomes were development of any malignancy and individual cancers as identified by ICD-9/ICD-10 codes. Results were adjusted for age at index date, index year, nulliparity, race, smoking, obesity, number of visits per year and highest level of education.

MAIN RESULTS AND THE ROLE OF CHANCE: Infertile women had an overall higher risk of developing cancer compared to non-infertile women (2.0 versus 1.7%, adjusted hazard ratio (aHR) = 1.18; CI: 1.12–1.24). In addition, the risk of uterine cancer (0.10 versus 0.06%, aHR = 1.78; CI: 1.39–2.28), ovarian cancer (0.14 versus 0.09%, aHR 1.64; CI: 1.33–2.01), lung cancer (0.21 versus 0.21%, aHR = 1.38; CI: 1.01–1.88), thyroid cancer (0.21 versus 0.16%, aHR = 1.29; CI: 1.09–1.53), leukemia (0.10 versus 0.06%, aHR = 1.55; CI: 1.21–1.98) and liver and gallbladder cancer (0.05 versus 0.03%, aHR = 1.59; CI: 1.11–2.30) were higher in infertile women compared to non-infertile women. In a subgroup analysis of women in each cohort who became pregnant and had a delivery during enrollment, the risk of uterine and ovarian cancer were similar between infertile and non-infertile women. In a subgroup analysis excluding women with PCOS and endometriosis from both cohorts, the risk of uterine cancer was similar between infertile and non-infertile women.

LIMITATIONS, REASONS FOR CAUTION: Absolute risk of cancer was low, average follow up for each individual was limited, and average age at index date was limited. Insurance databases have known limitations.

WIDER IMPLICATIONS OF THE FINDINGS: Using claims-based data, we report that infertile women may have a higher risk of certain cancers in the years after infertility evaluation; continued follow up should be considered after reproductive goals are achieved.

STUDY FUNDING/COMPETING INTEREST(S): None.

Key words: female infertility / fertility treatment / cancer / assisted reproduction / malignancy

Introduction

The association between infertility and the risk of developing a malignancy is an ongoing concern for both fertility patients and providers. Several studies suggest that infertile women are at higher risk of cancer than women from the general population, particularly hormone-sensitive cancers of the breast, ovaries and endometrium (Brinton *et al.*, 1989; Rossing *et al.*, 1994; Venn *et al.*, 1995; Meirow and

Schenker, 1996; Modan *et al.*, 1998; Althius *et al.*, 2005b; Rizzuto *et al.*, 2013; Practice Committee of ASRM, 2016; Skalkidou *et al.*, 2017; Williams *et al.*, 2018). There is some conflicting data, however, from additional studies showing no association between infertility and risk of breast and ovarian cancer (Venn *et al.*, 1995; Kashyap *et al.*, 2004; Jensen *et al.*, 2007; van den Belt-Dusebout *et al.*, 2016). Less is known about risks of non-gynecologic malignancies among infertile patients (Althius *et al.*, 2005b; Calderon-Margalit *et al.*, 2009). For all cancer types, the low incidence of malignancy in reproductive age women limits the conclusions of the majority of studies' risk associations. Furthermore, the effect of pregnancy on the risk association between infertility and cancer has been investigated for only a minority of cancer types (Nagasue *et al.*, 1986; Dupont and Page, 1987; Adami *et al.*, 1994; Lambe *et al.*, 1999; Kreuzer *et al.*, 2003; Giannitrapani *et al.*, 2006).

The objective of our study was to investigate whether infertility was associated with a subsequent risk of developing a malignancy. In addition, we sought to investigate the effect, if any, of pregnancy on the risk association between infertility and cancer. In order to overcome the low incidence of cancer in women of reproductive age, we used a large database of health insurance claims to examine outcomes in infertile patients as many insurance organizations offer coverage for infertility testing and treatment.

Materials and Methods

Patients

We analyzed subjects in the Optum® de-identified Clinformatics© Datamart between 2003 and 2016. Optum©'s Clinformatics® Data Mart (CDM) is derived from a database of administrative health claims for members of a large national managed care company affiliated with Optum. The database includes ~17–19 million annual covered lives, for a total of over 57 million unique lives over a 14-year period (1/2003 through 12/2016). These administrative claims submitted for payment by providers and pharmacies are verified, adjudicated, adjusted and de-identified prior to inclusion. The Clinformatics® Data Mart data comprises both commercial and Medicare Advantage health plan data. The population is geographically diverse, spanning all 50 states. This study was approved by the Stanford Institutional Review Board.

Given the variation in infertility coding and reimbursement practices in the USA, we attempted to be as broad as possible. The cohort of infertile women was comprised of women receiving any of the following: (i) an infertility diagnosis, (ii) fertility testing or (iii) fertility treatment. Women with an infertility diagnosis were identified by outpatient claims (ICD9: 628.x, 614.6, V26.89, ICD10: E23.0, N73.6, N97.x, Z31.81). Fertility testing was identified through diagnosis codes (V26.21, Z31.41) or the presence of a procedure code (CPT) for hysterosalpingogram (HSG) (74740). HSG was chosen to identify infertile women because it is commonly ordered as part of an initial fertility evaluation to assess tubal patency but is not otherwise part of routine gynecologic care (Crosignani and Rubin, 2000). Patients receiving fertility treatment were identified by the presence of a CPT code for intrauterine artificial insemination (58322), follicular puncture for oocyte retrieval (58970) or intrauterine embryo transfer (58974). The presence of a pharmacy claim for a prescription for clomiphene citrate or a gonadotropin (FSH, HMG, HCG) was also used to identify patients receiving fertility treatment. Clomiphene was chosen because it is routinely used to induce ovulation for women with anovulatory infertility as well as to promote superovulation for women who are ovulatory, and is not prescribed for non-fertility indications (Kousta *et al.*, 1997). Gonadotropins (FSH, HMG, HCG) have comprised the standard approach for ovulation stimulation and induction for ART since they were first implemented in 1981 and HCG has consistently been used to mimic the mid-cycle surge in luteinizing hormone that triggers final oocyte maturation in ART cycles (Lunenfeld, 2012).

The comparison group was composed of women receiving routine gynecologic care who did not have an infertility diagnosis or procedure codes for fertility testing or treatment. These patients were identified through the presence of a claim for a well woman visit (V72.31, Z01.411, Z01.419), encounter for contraceptive management (V25.0, V25.01–V25.04, V25.09, Z30.011, Z30.012, Z30.015, Z30.016, Z30.017, Z30.018, Z30.02, Z30.09), encounter for placement or removal of an IUD (V25.11–V25.13, Z30.430, Z30.432, Z30.433), encounter for placement of a contraceptive implant (V25.5, Z30.8), encounter for bilateral tubal ligation (V25.2, Z30.2), encounter for contraceptive surveillance (V25.40–V25.43, V25.49, Z30.40, Z30.41, Z30.42, Z30.431, Z30.436, Z30.44, Z30.45, Z30.49) and encounter for pap smear (V72.32, Z12.4, 88141–88155, 88164–88167, 88174-5, Q0091, G0101).

In both groups, women who became pregnant and had a delivery were identified by diagnosis and procedure codes indicating the end of a pregnancy. These diagnosis and procedure codes were obtained from a literature search of insurance claims data used to identify various pregnancy outcomes (Bennett *et al.*, 2014; Ailes *et al.*, 2016) and are listed in Supplementary Table SI.

We recorded the first date of a relevant diagnosis or procedure code as the index date. For patients in the infertile group, index date was the date of infertility diagnosis, testing or treatment. For patients in the non-infertile group, index date was the date of encounter for any of the services listed above. In order to be included in the study, patients were required to be enrolled in a plan covered by the database for at least 6 months before and after the index date. Patients were also required to be between 20 and 45 years old on the index date. In all groups, patients with a prior cancer diagnosis or with a cancer diagnosis within the 6 months following the index date were excluded from the study. This was identified through the presence of any claim with a diagnosis code for cancer.

Outcome ascertainment

Cancer diagnoses were identified using diagnosis codes on inpatient and outpatient claims. Diagnosis codes to identify cancer were aligned to the Surveillance, Epidemiology, and End Results (SEER) definitions. We identified patients with claims diagnoses indicating the presence of any invasive cancer excluding non-invasive cancers, skin squamous cell and skin basal cell cancers (ICD9: 140–209, C00.x-C96.x). We also identified the presence of specific cancers including upper respiratory (140.x-149.x, 160.x, 161.x, C00.x-C14.x, C30.x, C31.x, C32.x), stomach (151.x, C16.x), colorectal (153.x, 154.0, 154.1, 154.8, C18.x, C218.x, C19.x, C20.x), liver and gallbladder (155.x, 156.x, C22.x, C23.x), pancreas (157.x, C25.x), lung (162.x, C33.x, C34.x), melanoma (172.x, C43.x, D03.x), breast (174.x, C50.019, C50.119, C50.219, C50.319, C50.419, C50.519. C50.619, C50.819, C50.919), bladder (188.x, C67.x), kidney (189.0, 189.1, C64.9x, C65.9x), brain and nervous system (191.x, 192.x, C71.x, C72.50x, C72.1x, C70.0x, C70.1x, C70.9x, C72.9x, C72.0x), thyroid (193.x, C73.x), non-Hodgkin lymphoma (200.x, 202.x), Hodgkin lymphoma (201.x, C81.x), leukemia (204.x, 205.x, 206.x, 207.x, 208.x, C91.x, C92.x, C93.x, C94.x. C95.x), esophagus (150.x, C15.x), uterus (179.x, 182.x, C55.x, C54.0x, C54.1x, C54.2x, C54.3x, C54.8x, C54.9x), cervix (180.x, C53.0x, C53.1x, C53.8x, C53.9x), and ovary (183.x, C56.9x, C57.3x, C57.4x, C57.00x, C57.10x, C57.20x). Outcomes with <11 absolute number of events are required to be reported as <11 to protect patient privacy per the data usage agreement with Optum.

Several subset analyses were performed as follows: including only women in both infertile and non-infertile cohorts who became pregnant and had a delivery during the enrollment period, excluding women with PCOS (identified with diagnosis codes 256.4, 628.0, N97.0, E28.2) and endometriosis (identified with diagnosis codes 617.0, N80.9) from both infertile and non-infertile cohorts, and limiting follow up of all women to >2, >3 and >4 years. For all subset analyses, the incidence of a cancer diagnosis was compared between infertile and non-infertile cohorts.

Confounder selection

Cancer risk has been attributed in the literature to a variety of demographic factors including age, year of presentation, reproductive history, access to care, race, smoking, obesity and highest level of education (American Cancer Society, 2019). Detection of cancer is also influenced by screening practices including screening mammography (Tabar *et al.*, 1985). In the Optum database, several demographic factors are available including age, year of diagnosis, race and highest level of education are available and were recorded at the index date. For each patient, diagnosis codes were then used to identify obesity (278.0, E66.9, E66.01, E66.3, E66.2), smoking (305.1, V15.82, F17.200, Z87.891), and nulliparity (V22.0, V23.81, V23.83, O0.95, O0.96). Diagnosis codes entered either at the index date or during the follow up period were included. As a proxy measure of access to care, for each patient, the number of outpatient visits after the index date was determined based on the presence of claims for CPT codes indicating new and follow up visits, consultations, or preventive medicine encounters. For patients with a diagnosis of breast cancer, those who obtained a screening mammogram prior to cancer diagnosis were identified using codes (ICD-10: Z12.31, ICD-9: V76.11, CPT: 77067, 77066, 77065, 77063, HCPCS: G0202, G0204, G0206).

Statistical analysis

To compare demographics between infertile and non-infertile cohorts, the Chi-square test was used for categorical variables, Student's T-test was used to compare age, and the Wilcoxon rank-sum test was used to compare follow up time and number of visits. Patients accrued at risk time beginning from their index dates until cancer diagnosis or the last enrollment date in a health plan in the Optum® insurance claims database. The risk of cancer between infertile patients and the non-infertile group was assessed using a Cox proportional hazards model while adjusting for age at index date, index year, nulliparity, race, smoking, obesity, number of visits per year and highest level of education. Risk of breast cancer was also adjusted for receiving a screening mammogram. All p values were 2-sided with $P < 0.05$ considered statistically significant. Analyses were performed using SAS (version 9.4, SAS Institute, Inc., Cary, NC, USA).

Results

Patient demographics: infertile versus non-infertile patients

Overall, 64 345 patients had a diagnosis of infertility, underwent fertility testing or underwent fertility treatment. The non-infertile group was comprised of 3 128 345 patients who were deemed non-infertile. Patients in the infertile group were older on average at the index date (34.0 ± 5.7 years) compared to the non-infertile group (32.7 ± 7.4 years) ($P < 0.0001$). Patients were, on average, followed for 3.8 ± 3.3 years in the infertile group and 3.9 ± 3.3 years in the non-infertile group. In the study interval, there were 246 485 person-years of follow up in the infertile group and 12 268 968 person-years of follow up in the non-infertile group. Infertile patients were more likely to be nulliparous, obese and smokers compared to non-infertile patients ($P < 0.0001$ for all comparisons) (Table I). Infertile patients had a higher median number of visits per person per year (4.1 visits) compared to the non-infertile group (2.7 visits) ($P < 0.0001$). In both groups, the majority of patients were Caucasian. Education level, income and geographic distribution were similar between the infertile and non-infertile groups. Comparison of all demographics between infertile and non-infertile cohorts was significant ($P < 0.001$), likely due to the large sample size.

Patient demographics: subset of patients with a delivery

Of the 64 345 patients in the infertile group, 22 024 (34.2%) had a pregnancy and subsequent delivery during the enrollment period (4.5 ± 3.3 years). Of the 3 128 345 patients in the non-infertile group, 626 532 (20.0%) had a pregnancy and subsequent delivery during the enrollment period (4.5 ± 3.4 years). A higher proportion of women became pregnant and delivered in the infertile cohort compared to the non-infertile cohort. Infertile patients were older on average at time of follow up (32.6 ± 4.9 years) compared to non-infertile patients (29.8 ± 5.2 years) (Table I). Infertile patients in the delivery subgroup were more likely to be nulliparous and obese and less likely to be smokers compared to non-infertile patients.

Overall incidence of cancer across study cohorts

During follow up, there were 1310 cancer diagnoses in the infertile group and 53 116 cancer diagnoses in the non-infertile group. In both groups, the most common cancer diagnosed was breast (329 cases in the infertile group and 15 348 cases in the non-infertile group) (Table II).

Comparison of infertile and non-infertile cohorts

Women with infertility had an overall higher risk of developing cancer compared to non-infertile women (2.0 versus 1.7%, adjusted hazard ratio (aHR) = 1.18; CI: 1.12–1.24). In addition, the risk of uterine cancer (0.10 versus 0.06%, aHR = 1.78; CI: 1.39–2.28), ovarian cancer (0.14 versus 0.09%, aHR = 1.64; CI: 1.33–2.01), lung cancer (0.21 versus 0.21%, aHR = 1.38; CI: 1.01–1.88), thyroid cancer (0.21 versus 0.16%, aHR = 1.29; CI: 1.09–1.53), leukemia (0.10 versus 0.06%, aHR = 1.55; CI: 1.21–1.98) and liver and gallbladder cancer (0.05 versus 0.03%, aHR = 1.59; CI: 1.11–2.30) were higher in infertile women compared to non-infertile women. The risk of breast cancer was similar between the infertile and non-infertile groups (0.51 versus 0.49%, aHR = 1.08; CI: 0.97–1.21). Adjusting for covariates including age at index date, index year, nulliparity, race, smoking, obesity, number of visits per year and education did not significantly affect any of the results. Adjusting the risk of developing breast cancer for receiving a screening mammogram did not significantly affect the results. A sensitivity analysis was performed excluding obesity and smoking as potential confounders (Supplementary Table SII) with no significant change in the point estimates.

Table I Patient demographics of overall infertile cohort and subset of patients within each cohort with pregnancy and subsequent delivery during follow up period.

	Full cohort		Delivery cohort	
	Infertile	Control	Infertile	Control
Number of patients	64 345	3 128 345	22 024	626 532
Age at index date (years)				
Mean (SD)	34.0 (5.7)	32.7 (7.4)	32.6 (4.9)	29.8 (5.2)
20–24	3382 (5.3)	555 690 (17.8)	1088 (4.9)	104 136 (16.6)
25–29	11 669 (18.1)	625 505 (20.0)	4963 (22.5)	207 144 (33.1)
30–34	18 962 (29.5)	604 782 (19.3)	8100 (36.8)	196 510 (31.4)
35–39	18 102 (28.1)	592 800 (19.0)	5980 (27.2)	95 339 (15.2)
40–45	12 230 (19.0)	749 568 (24.0)	1893 (8.6)	23 403 (3.7)
Follow up time (years)[b]				
Mean (SD)	3.8 (3.3)	3.9 (3.3)	4.5 (3.3)	4.5 (3.4)
0–1	10 760 (16.7)	496 814 (15.9)	1261 (5.7)	69 482 (11.1)
1–2	15 401 (23.9)	717 058 (22.9)	4879 (22.2)	122 604 (19.6)
2–3	9246 (14.4)	454 847 (14.5)	3657 (16.6)	89 904 (14.4)
3–4	6346 (9.9)	318 191 (10.2)	2700 (12.3)	69 284 (11.1)
4+	22 592 (35.1)	1 141 435 (36.5)	9527 (43.3)	275 258 (43.9)
Total	246 484.6	12 268 968.2	98 145.3	2 791 449.2
Nulliparity	11 396 (17.7)	266 843 (8.5)	9421 (42.8)	232 004 (37.0)
Obesity	11 826 (18.4)	425 310 (13.6)	3586 (16.3)	81 826 (13.1)
Smoking	7182 (11.2)	303 955 (9.7)	1559 (7.1)	53 654 (8.6)
Index date				
2003	5903 (9.2)	309 940 (9.9)	2179 (9.9)	63 579 (10.2)
2004	7481 (11.6)	334 317 (10.7)	2644 (12.0)	68 067 (10.9)
2005	6154 (9.6)	309 144 (9.9)	2135 (9.7)	62 792 (10.0)
2006	5181 (8.1)	278 365 (8.9)	1777 (8.1)	57 446 (9.2)
2007	4928 (7.7)	261 200 (8.4)	1736 (7.9)	54 314 (8.7)
2008	4675 (7.3)	244 621 (7.8)	1640 (7.5)	52 091 (8.3)
2009	4234 (6.5)	220 298 (7.0)	1552 (7.1)	45 982 (7.3)
2010	3581 (5.6)	184 858 (5.9)	1323 (6.0)	38 196 (6.1)
2011	3702 (5.8)	187 229 (6.0)	1358 (6.2)	38 148 (6.1)
2012	3759 (5.8)	18 1671 (5.8)	1356 (6.2)	36 771 (5.9)
2013	3934 (6.1)	174 377 (5.6)	1368 (6.2)	34 404 (5.5)
2014	3994 (6.2)	168 622 (5.4)	1407 (6.4)	32 262 (5.2)
2015	4367 (6.8)	18 0157 (5.8)	1252 (5.7)	30 109 (4.8)
2016	2452 (3.8)	93 546 (3.0)	297 (1.4)	12 371 (2.0)
Visits per person year				
Median (range)	4.10 (0–92.5)	2.70 (0–236.2)	4.50 (0–61.0)	2.6 (0–236.2)
<1	7973 (12.4)	581 663 (18.6)	1738 (7.9)	111 038 (17.7)
1–2	8209 (12.8)	60 9041 (19.5)	2452 (11.1)	129 154 (20.6)
2+	48 163 (74.9)	193 7641 (61.9)	17 834 (81.0)	386 340 (61.7)
Race/Ethnicity				
White	39 492 (61.4)	2 140 071 (68.4)	14 123 (64.1)	409 673 (65.4)
Asian	6802 (10.6)	165 505 (5.3)	2871 (13.0)	47 970 (7.7)
Black	6610 (10.3)	319 552 (10.2)	1552 (7.1)	57 303 (9.2)
Hispanic	7966 (12.4)	363 712 (11.6)	2220 (10.1)	81 712 (13.0)
Unknown	3475 (5.4)	139 505 (4.5)	1258 (5.7)	29 874 (4.8)

Continued

Table I *Continued*

	Full cohort		Delivery cohort	
	Infertile	Control	Infertile	Control
Education				
Less than 12th grade	520 (0.8)	21 867 (0.7)	84 (0.4)	4766 (0.8)
High School Diploma	14 942 (23.2)	761 574 (24.3)	3746 (17.0)	146 114 (23.3)
Less than Bachelor degree	32 641 (50.7)	1 651 487 (52.8)	10 935 (49.7)	334 771 (53.4)
Bachelor Degree plus	15 868 (24.7)	676 596 (21.6)	7151 (32.5)	137 753 (22.0)
Unknown	374 (0.6)	16 821 (0.5)	108 (0.5)	3128 (0.5)
Income				
<50 K	8983 (14.0)	486 152 (15.5)	2235 (10.2)	93 047 (14.9)
50–100 K	14 237 (22.1)	713 358 (22.8)	4856 (22.1)	148 950 (23.8)
100 K+	17 184 (26.7)	819 841 (26.2)	7765 (35.3)	168 350 (26.9)
Unknown	23 941 (37.2)	1 108 994 (35.5)	7168 (32.6)	216 185 (34.5)
Region of the country				
Midwest	15 538 (24.2)	814 816 (26.1)	5467 (24.8)	165 330 (26.4)
Northeast	8285 (12.9)	324 942 (10.4)	3304 (15.0)	61 525 (9.8)
South	26 123 (40.6)	1 426 092 (45.6)	8155 (37.0)	275 633 (44.0)
West	14 187 (22.1)	557 727 (17.8)	5041 (22.9)	123 053 (19.6)
Unknown	212 (0.3)	4768 (0.2)	57 (0.3)	991 (0.2)

Unless stated otherwise all data are *n* (%).
[a]Comparison of all demographics between infertile and non-infertile cohorts was significant ($P < 0.001$) likely due to the large sample size.
[b]Follow-up time was calculated from the index date to the last enrolled date in the database.

Table II Absolute incidence, events per person years, hazard ratios and 95% CI for the association between female infertility and incidence of cancer.

	Infertile (246 484 person years)		Control (12 268 968 person years)		Hazard ratio[a] (95% CI)
	n (%)	Events/1 K person years	*n* (%)	Events/1 K person years	Infertile versus control
All cancer diagnoses	1310 (2.04)	5.31	53 116 (1.70)	4.33	1.18 (1.12–1.24)
Breast	329 (0.51)	1.33	15 348 (0.49)	1.25	1.08 (0.97–1.21)[b]
Uterus	66 (0.10)	0.27	1779 (0.06)	0.14	1.78 (1.39–2.28)
Cervix	54 (0.08)	0.22	2400 (0.08)	0.20	1.06 (0.81–1.38)
Ovary	93 (0.14)	0.38	2705 (0.09)	0.22	1.64 (1.33–2.01)
Urinary bladder	19 (0.03)	0.08	662 (0.02)	0.05	1.38 (0.87–2.18)
Kidney	19 (0.03)	0.08	972 (0.03)	0.08	0.93 (0.59–1.46)
Lung	133 (0.21)	0.54	6453 (0.21)	0.53	1.38 (1.01–1.88)
Melanoma	133 (0.21)	0.54	6453 (0.21)	0.53	0.996 (0.84–1.18)
Thyroid	138 (0.21)	0.56	4869 (0.16)	0.40	1.29 (1.09–1.53)
Non-Hodgkins Lymphoma	67 (0.10)	0.27	3331 (0.11)	0.27	0.93 (0.73–1.19)
Hodgkins Lymphoma	18 (0.03)	0.07	846 (0.03)	0.07	0.99 (0.62–1.58)
Leukemia	66 (0.1)	0.27	1987 (0.06)	0.16	1.55 (1.21–1.98)
Brain/CNS	43 (0.07)	0.17	1886 (0.06)	0.15	1.08 (0.80–1.47)
Upper digestive tract	35 (0.05)	0.14	1620 (0.05)	0.13	1.02 (0.73–1.43)
Colon and rectum	61 (0.09)	0.25	2621 (0.08)	0.21	1.15 (0.89–1.48)
Liver and gallbladder	30 (0.05)	0.12	889 (0.03)	0.07	1.59 (1.11–2.3)
Pancreas	11 (0.02)	0.04	560 (0.02)	0.05	0.94 (0.52–1.72)

[a]Adjusted for age at index date, index year, nulliparity, race, smoking, obesity, number of visits per year and highest level of education.
[b]Additionally adjusted for screening mammogram.

Subset analysis of patients with a delivery during enrollment

A subset analysis of cancer incidence in women who became pregnant and had a delivery in both the infertile and non-infertile groups was performed. The overall risk of cancer was higher in the infertile subgroup (1.78 versus 1.16%, aHR = 1.18; CI: 1.06–1.30), however, some associations were lower in magnitude (Table III). Specifically, the risk of uterine (0.04 versus 0.02%, aHR = 1.31; CI: 0.64–2.69), ovarian cancer (0.09 versus 0.05%, aHR = 1.41; CI: 0.88–2.25), lung (0.04 versus 0.03%, aHR = 1.18; CI: 0.60–2.31) and liver and gallbladder cancer (0.04 versus 0.02%, aHR = 1.51; CI: 0.73–3.12) were similar between the infertile and non-infertile subgroups. No new associations between infertility and cancer risk were noted in this subset analysis. Adjusting for the above mentioned covariates did not significantly affect the results.

Subset analysis excluding patients with PCOS and endometriosis from both cohorts

In the infertile cohort, 36.3% of patients had a diagnosis of PCOS or endometriosis. In the non-infertile group, 4.0% of patients had a diagnosis of PCOS or endometriosis. A subset analysis was performed excluding patients with PCOS and endometriosis from both infertile and non-infertile cohorts. The overall risk of cancer was higher in the infertile subgroup (1.91 versus 1.63%, aHR = 1.14; CI: 1.06–1.22). Risk associations of individual cancers followed the same trends as the overall analysis with the exception of uterine cancer which did not differ between this infertile and non-infertile subgroup (0.05 versus 0.05%, aHR = 0.99; CI: 0.64–1.55).

Subset analysis limiting follow up time in both cohorts

The mean follow up time for both infertile and non-infertile cohorts was 3.9 years. A subset analysis was performed limiting follow up time to >2, >3 and >4 years in both infertile and non-infertile cohorts. The overall risk of any malignancy and risks of individual cancers was unchanged between the infertile and non-infertile subgroups compared to the overall analysis (data not shown).

Discussion

Although the overall incidence of cancer is low for all women in the cohort, we identify an association between infertility and higher risk of cancer compared to a group of non-infertile women.

Comparison of infertile and non-infertile cohorts

Higher risk of cancer in infertile women compared to the general population has been demonstrated in several prior studies. In particular, the risk of developing hormone-sensitive cancers of the breast, ovaries and uterus has previously been reported in infertile patients and attributed to anovulation and dysregulation of estradiol and progesterone levels (Brinton *et al.*, 1989; Rossing *et al.*, 1994; Venn *et al.*, 1995; Meirow and Schenker, 1996; Modan *et al.*, 1998). During fertility treatment, estradiol and progesterone levels are upregulated above physiologic levels (Joo *et al.*, 2010); therefore, concern has also been raised regarding fertility treatment and the risk of developing hormone-sensitive cancers. For example, the incidence of uterine cancer was found to be higher after clomiphene treatment, with the strongest correlation in obese and nulliparous women (Althius *et al.*, 2005a).

In both infertile and non-infertile patients, we report a low, overall absolute risk of cancer (2.04 and 1.70%, respectively). Comparison of cancer risk between infertile and non-infertile groups reveals an 18% higher risk of cancer among infertile patients compared to non-infertile patients. The low overall incidence, however, translates into a very modest increase in absolute risk of 1/49 compared to 1/59 for infertile and non-infertile patients, respectively. Given the overall risk association, we then sought to identify individual cancers risk associations. Among infertile patients, we report a higher risk of ovarian and uterine cancers and a similar risk of breast cancer compared to non-infertile patients. Furthermore, we report a higher risk of non-hormonal cancers including lung cancer, thyroid cancer, liver and gallbladder cancer and leukemia in the infertile group compared to non-infertile patients. While several associations were significant, absolute increases in risk were modest due to the low overall incidence of cancer in our patient cohort.

Subset analyses

Cancers typically develop over several years, and occult carcinomas may impair fertility for several years before clinical detection (Levanon *et al.*, 2008). Cancer risk has also been widely associated with several demographic factors (American Cancer Society, 2019). Infertile patients had higher rates of obesity and smoking compared to non-infertile patients, both of which can increase the risk of cancer (Bianchini *et al.*, 2002; Bach *et al.*, 2003) and therefore were controlled for in our statistical analysis. A secondary analysis in which obesity and smoking were excluded as potential confounders resulted in similar point estimates to the overall analysis. Infertile patients were, on average, older than non-infertile patients at the index date, however, the mean age difference of 1.3 years is unlikely to be clinically significant. Patients in the infertile group did have a higher number of visits per year compared to the non-infertile group, and perhaps were more likely to report symptoms that eventually led to a cancer diagnosis. Infertile patients may also have additional comorbidities that were not controlled for in the analysis and may result in higher risk of cancer. For example, PCOS has been consistently associated with increased risk of uterine cancer and sporadically associated with increased risk of breast and ovarian cancer (Balen, 2001; Dumesic and Lobo, 2013; Barry *et al.*, 2014). Endometriosis has been associated with increased risk of ovarian cancer (Somigliana *et al.*, 2006). In a subgroup analysis excluding patients with PCOS and endometriosis, we identified similar point estimates for the overall risk of malignancy and for individual cancers with the exception of uterine cancer. This loss of a risk association with uterine cancer suggests that PCOS/endometriosis may confound the relationship between infertility and uterine cancer. Another potential common mechanism between cancer and infertility is the presence of mutations in genes associated with DNA repair (e.g. BRCA) that may increase the risk of both conditions (Giordano *et al.*, 2016; Johnson *et al.*, 2017). In summary, our findings suggest an association between infertility and higher risk of cancer compared to a population deemed by claims-based criteria to be non-infertile.

Table III Absolute incidence, events per person years, hazard ratios and 95% CI for the association between female infertility and incidence of cancer in the subset of patients with pregnancy and subsequent delivery during follow up.

	Infertile subset (98 036 person years)		Control subset (2 787 915 person years)		Hazard ratio[a] (95% CI)
	n (%)	Events/1 K person years	*n* (%)	Events/1 K person years	Infertile versus control
All cancer diagnoses	393 (1.78)	4.00	7275 (1.16)	2.61	1.18 (1.06–1.30)
Breast	102 (0.46)	1.04	1876 (0.30)	0.67	1.03 (0.85–1.26)[b]
Uterus[c]	<11	0.08	135 (0.02)	0.05	1.31 (0.64–2.69)
Cervix	12 (0.05)	0.12	342 (0.05)	0.12	0.90 (0.51–1.61)
Ovary	19 (0.09)	0.19	295 (0.05)	0.11	1.41 (0.88–2.25)
Urinary bladder[c]	<11	0.04	71 (0.01)	0.03	1.12 (0.41–3.09)
Kidney[c]	<11	0.04	111 (0.02)	0.04	0.81 (0.30–2.20)
Lung[c]	<11	0.09	172 (0.03)	0.06	1.18 (0.60–2.31)
Melanoma	45 (0.20)	0.46	1049 (0.17)	0.38	0.98 (0.72–1.32)
Thyroid	57 (0.26)	0.58	906 (0.14)	0.32	1.40 (1.07–1.84)
Non-Hodgkins Lymphoma	19 (0.09)	0.19	455 (0.07)	0.16	0.94 (0.59–1.50)
Hodgkins Lymphoma[c]	<11	0.08	116 (0.02)	0.04	1.75 (0.85–3.61)
Leukemia	19 (0.09)	0.19	257 (0.04)	0.09	1.78 (1.11–2.85)
Brain/CNS	11 (0.05)	0.11	257 (0.04)	0.09	0.97 (0.53–1.78)
Upper digestive tract[c]	<11	0.04	228 (0.04)	0.08	0.40 (0.15–1.09)
Colon and rectum	13 (0.06)	0.13	327 (0.05)	0.12	0.83 (0.48–1.46)
Liver and gallbladder[c]	<11	0.08	108 (0.02)	0.04	1.51 (0.73–3.12)
Pancreas[c]	<11	0.02	61 (0.01)	0.02	0.75 (0.18–3.09)

[a]Adjusted for age at index date, index year, nulliparity, race, smoking, obesity, number of visits per year and highest level of education.
[b]Additionally adjusted for screening mammogram.
[c]Absolute number of events not reported. Outcome instead reported as <11 to protect patient privacy per data usage agreement.

In a subset analysis, we compared outcomes of women in each cohort who became pregnant and had a delivery during the enrollment period. A higher proportion of the infertile cohort underwent pregnancy and delivery during the enrollment period than the non-infertile cohort. This is likely multifactorial in nature and may be due to non-infertile women already completing their family prior to the index date and/or not attempting to become pregnant during the enrollment period. In addition, a higher proportion of both the infertile cohort and non-infertile cohort who had a child during the enrollment period were nulliparous compared to the overall cohort. This subset of patients were younger than the overall cohort and likely had a higher prevalence of primary infertility. In patients who became pregnant and had a delivery, several risk associations that were noted in the overall comparison were diminished. Namely, the risk of developing ovarian, uterine, lung, liver and gallbladder cancer were similar between the subset of infertile and non-infertile patients who became pregnant with a subsequent delivery but were elevated in the overall comparison of infertile women to non-infertile women. As some of these cancers are hormone mediated, it does suggest a possible mechanism behind the seemingly protective effect of childbearing among infertile women who succeed in fertility treatment. It is important to note that cancers in reproductive age women are, overall, rare events and the conclusions of our subgroup analysis are limited by the low incidence of the outcomes of interest. Our findings are in accordance with previous reports that parity is associated with a reduction in the risk of ovarian and uterine cancer (Adami *et al.*, 1994; Lambe *et al.*, 1999). Pregnancy resulting in delivery has also been shown to be protective against breast cancer (Dupont and Page, 1987). We report no association between breast cancer risk in either the overall infertile cohort or the subset of infertile patients with a delivery. The effect of pregnancy on non-hormone sensitive cancers is not as clearly established. No clear association has been reported between prior live births and lung cancer risk (Kreuzer *et al.*, 2003). The role of estrogen in the development of hepatic adenomas has been well established and by extension the role of reproductive hormones in development of liver cancer has been hypothesized but not clearly demonstrated (Nagasue *et al.*, 1986; Giannitrapani *et al.*, 2006). In order to elucidate if infertility treatment or childbearing are responsible for reducing risk associations with cancer, future studies would examine outcomes of infertile patients who conceive spontaneously versus with fertility treatment.

Study limitations

While the number of patients and total person-years of follow up was high, one of the limitations of the study is the low incidence of cancer outcomes in reproductive age women. Furthermore, population health level databases are subject to a variety of limitations. While the number of total person years of follow up is large, the average number of years of follow up for each patient and their age at index date were limited. It is of note, however, that several significant risk associations were noted in a short duration of follow up. It is not known whether factors related to population turnover in the database are related to

cancer risk, and thus is a source of potential bias. We did perform a sensitivity analysis limited to women with progressively longer follow up to determine how shorter follow up affected results. We identified similar point estimates in these analyses. Furthermore, the limited follow up time in our study inadvertently selects for the subset of cancers that present at an early age. Future studies would incorporate longer term follow up after infertility diagnosis or initiation of fertility treatment in order to more effectively capture these rare outcomes. While we controlled for several covariates in our data analysis, our ability to capture some was limited. For example, nulliparous status is not consistently identified by insurance claims data. In addition, we could not distinguish between a history of smoking and current smoking status. While the Optum® Data Mart data comprises both commercial and Medicare Advantage health plan data from a geographically diverse population, there is selection bias of the population as a whole as they are all insured. The selection of patients for the infertile and non-infertile cohorts is subject to some limitations. Estimates of the prevalence of infertility are around 10% (Chandra *et al.*, 2013), while in our study population ~2% of women were infertile. Infertility may have been undercounted, for example, among nulliparous women in the non-infertile cohort who were utilizing contraception and therefore do not have proven fertility. In addition, while the majority of inseminations included are likely utilizing partner sperm, we are unable to distinguish these from donor insemination cycles in which only a portion of women are infertile. We expect, however, that misclassification would be non-differential which would shift results towards the null so any association we identify is likely to be underestimated. Future studies would separately examine the incidence of cancer in women diagnosed with infertility compared to infertile women who subsequently undergo fertility treatment. Details regarding infertility diagnosis and treatment are not fully captured in insurance claims data, therefore we did not separately analyze risk associations with infertility diagnosis versus treatment. For example, patients could have sought fertility treatment outside of insurance coverage. This would have not been captured in our database and led to non-differential misclassification with a regression to a null finding. Despite this likely underestimate of associations, we still found differences between the overall infertile and non-infertile cohorts. While we included women utilizing contraception in the non-infertile cohort, we did not account for potential risk modification associated with IUD and oral contraceptive use (Cramer, 2012). Future studies would control for potential confounders including oral and long-acting contraceptive use, BRCA diagnosis, age at first birth, and oligomenorrhea, which may affect both fertility and cancer risk. As the database we utilized was de-identified, linkage to national registries and death certificates was not possible. While we did include inpatient and outpatient claims data, care received outside the insurance system or care that was incorrectly coded would be missed. The limitations of claims-based data have been previously described (Roos *et al.*, 1993); in the absence of linkage to a population-based registry, we are unable to distinguish subject misidentification or loss of continuity. For example, we were not able to separately identify women diagnosed with infertility who then conceive spontaneously. This would require linking ART treatment records to birth certificates, a level of granularity that we were not able to access in the de-identified database, and thus is a potential source of detection bias. Finally, due to the high number of cancer outcomes investigated, there is a possibility that a small fraction of the results are false positives due to the role of chance. We investigated 17 individual cancers and found 6 positive associations. Chance alone may account for one false positive but would be unlikely to explain all of them.

Using claims-based data, we report that infertility is associated with a higher risk of cancer compared to a group of non-infertile women. From a subset analysis, we report that infertile women who conceive may have a lower risk of certain malignancies, namely uterine and ovarian cancers. While the absolute increase in cancer risk with infertility is small, this increase was seen within only 4 years of infertility diagnosis, strongly supporting the need for further study to determine what factors influence the long-term cancer risk for infertile women.

Supplementary data

Supplementary data are available at *Human Reproduction* online.

Authors' roles

G.M., R.B.L., V.L.B. and M.L.E. participated in study design and article preparation. S.L. participated in study design and statistical analysis. All authors participated in reviewing and approving the final version of this article.

Funding

None.

Conflict of interest

None.

References

Adami HO, Lambe M, Persson I, Ekbom A, Adami HO, Hsieh CC, Trichopoulos D, Ekbom A, Lambe M, Leon D *et al*. Parity, age at first childbirth and risk of ovarian cancer. *Lancet* 1994;**344**:1250–1254.

Ailes EC, Simeone RM, Dawson AL, Peterson EE, Gilboa SM. Using insurance claims data to identify and estimate critical periods in pregnancy: an application to antidepressants. *Birth Defects Res A Clin Mol Teratol* 2016;**106**:927–934.

Althius MD, Moghissi KS, Westhoff CL, Scoccia B, Lamb EJ, Lubin JH, Brinton LA. Uterine cancer after use of clomiphene citrate to induce ovulation. *Am J Epidemiol* 2005a;**161**:607–615.

Althuis MD, Scoccia B, Lamb EJ, Moghissi KS, Westhoff CL, Mabie JE, Brinton LA. Melanoma, thyroid, cervical, and colon cancer risk after use of fertility drugs. *Am J Obstet Gynecol* 2005b;**193**:668–674.

American Cancer Society. *Cancer Facts and Figures 2019*. Atlanta: American Cancer Society, 2019.

Bach PB, Kattan MW, Thornquist MD, Gris MG, Tate RC, Barnett MJ, Hsieh LJ, Begg CB. Variations in lung cancer risk among smokers. *J Natl Cancer Inst* 2003;**95**:470–478.

Balen A. Polycystic ovarian syndrome and cancer. *Hum Reprod Update* 2001;**7**:522–525.

Barry JA, Azizia MM, Hardiman PJ. Risk of endometrial, ovarian, and breast cancer in women with polycystic ovarian syndrome: a systematic review and meta-analysis. *Hum Reprod Update* 2014;**20**:748–758.

Bennett WL, Chang HY, Levine DM, Wang L, Neale D, Werner EF, Clark JM. Utilization of primary and obstetric care after medically complicated

pregnancies: an analysis of medical claims data. *J Gen Intern Med* 2014; **29**:636–645.

Bianchini F, Kaaks R, Vainio H. Overweight, obesity and cancer risk. *Lancet Oncol* 2002;**3**:565–574.

Brinton LA, Melton LJ III, Malkasian GD Jr, Bond A, Hoover R. Cancer risk after evaluation for infertility. *Am J Epidemiol* 1989;**129**:712–722.

Calderon-Margalit R, Friedlander Y, Yanetz R, Kleinhaus K, Perrin MC, Manor O, Harlap S, Paltiel O. Cancer risk after exposure to treatments for ovulation induction. *Am J Epidemiol* 2009;**169**:365–375.

Chandra A, Copen C, Stephen EH. Infertility and impaired fecundity in the United States, 1982-2010: data from the national survey of family growth. *Nat Health Stat Rep* 2013; no. 67. https://pdfs.semanticscholar.org/1cca/08d799cca2a803e2884291600a8a88ca933c.pdf.

Cramer DW. The epidemiology of endometrial and ovarian cancer. *Hematol Oncol Clin North Am* 2012;**26**:1–12.

Crosignani PG, Rubin PL. Optimal use of infertility diagnostic tests and treatments. The ESHRE Capri Workshop Group. *Hum Reprod* 2000;**15**:723–732.

Dumesic DA, Lobo RA. Cancer risk and PCOS. *Steroids* 2013;**78**:782–785.

Dupont WD, Page DL. Breast cancer risk associated with proliferative disease, age at first birth and a family history of breast cancer. *Am J Epidemiol* 1987;**125**:769–779.

Giannitrapani L, Soresi M, La Spada E, Cervello M, D'Alessandro N, Montalto G. Sex hormones and risk of liver tumor. *Ann N Y Acad Sci* 2006;**1089**:228–236.

Giordano S, Garrett-Mayer E, Mittal N, Smith K, Shulman L, Passaglia C, Gradishar W, Pavone ME. Association of BRCA1 mutations with impaired ovarian reserve: connection between infertility and breast/ovarian cancer risk. *J Adolesc Young Adult Oncol* 2016;**5**:337–343.

Jensen A, Sharif H, Svare EI, Frederiksen K, Kjaer SK. Risk of breast cancer after exposure to fertility drugs: results from a large Danish cohort study. *Cancer Epidemiol Biomarkers Prev* 2007;**16**:1400–1407.

Johnson L, Sammel MD, Domchek S, Schanne A, Prewitt M, Gracia C. Antimullerian hormone levels are lower in BRCA2 mutation carriers. *Fertil Steril* 2017;**107**:1256–1265.

Joo BS, Park SH, An BM, Kim KS, Moon SE, Moon HS. Serum estradiol levels during controlled ovarian hyperstimulation influence the pregnancy outcome of in vitro fertilization in a concentration-dependent manner. *Fertil Steril* 2010;**93**:442–446.

Kashyap S, Moher D, Fung MF, Rosenwaks Z. Assisted reproductive technology and the incidence of ovarian cancer: a meta-analysis. *Obstet Gynecol* 2004;**103**:785–794.

Kousta E, White DM, Franks S. Modern use of clomiphene citrate. *Hum Reprod Update* 1997;**3**:359–365.

Kreuzer M, Gerken M, Heinrich J, Kreienbrock L, Wichmann HE. Hormonal factors and risk of lung cancer among women. *Int J Epidemiol* 2003;**32**:263–271.

Lambe M, Wuu J, Weiderpass E, Hsieh CC. Childbearing at older age and endometrial cancer risk. *Cancer Causes Control* 1999;**10**:43–49.

Levanon K, Crum C, Drapkin R. New insights into the pathogenesis of serous ovarian cancer and its clinical impact. *J Clin Oncol* 2008;**26**:5284–5293.

Lunenfeld B. Gonadotropin stimulation: past, present and future. *Reprod Med Biol* 2012;**11**:11–25.

Meirow D, Schenker JG. The link between female infertility and cancer: epidemiology and possible aetiologies. *Hum Reprod Update* 1996;**2**:63–75.

Modan B, Ron E, Lerner-Geva L, Blumstein T, Menczer J, Rabinovici J, Oelsner G, Freedman L, Mashiach S, Lunenfeld B. Cancer incidence in a cohort of infertile women. *Am J Epidemiol* 1998;**147**:1038–1042.

Nagasue N, Yukaya H, Ogawa Y, Ito A. Estrogen receptors in hepatocellular carcinoma. *Cancer* 1986;**57**:87–91.

Practice Committee of the American Society for Reproductive Medicine. Fertility drugs and cancer: a guideline. *Fertil Steril* 2016;**106**:0015–0282.

Rizzuto I, Behrens RF, Smith LA. Risk of ovarian cancer in women treated with ovarian stimulating drugs for infertility. *Cochrane Database Syst Rev* 2013;**8**:CD008215.

Roos L, Mustard C, Nicol P, McLerran D, Malenka D, Young K, Cohen M. Registries and administrative data: organization and accuracy. *Med Care* 1993;**31**:201–212.

Rossing MA, Daling JR, Weiss NS, Moore DE, Self SG. Ovarian tumors in a cohort of infertile women. *N Engl J Med* 1994;**331**:771–776.

Skalkidou A, Sergentanis TN, Gialamas SP, Georgakis MK, Psaltopoulou T, Trivella M, Siristatidis C, Evangelou E, Petridou E. Risk of endometrial cancer in women treated with ovary-stimulating drugs for subfertility. *Cochrane Database Syst Rev* 2017;**3**:CD010931.

Somigliana E, Vigano P, Parazzini F, Stoppelli S, Giambattista E, Vercellini P. Association between endometriosis and cancer: a comprehensive review and critical analysis of clinical and epidemiological evidence. *Gynecol Oncol* 2006;**101**:331–341.

Tabar L, Gad A, Holmberg LH, Ljungquist U, Kopparberg County Project Group, Fagerberg CJG, Baldetorp L, Grontoft O, Lundstrom B, Manson JC *et al*. Reduction in mortality from breast cancer after mass screening with mammography: randomised trial from the breast cancer screening working group of the Swedish national board of health and welfare. *Lancet* 1985;**325**:829–832.

van den Belt-Dusebout AW, Spaan M, Lambalk CB, Kortman M, Laven JS, van Santbrink EJ, van der Westerlaken LJ, Cohlen BJ, Braat DDM, Smeenk JM *et al*. Ovarian stimulation for in vitro fertilization and long-term risk of breast cancer. *J Am Med Assoc* 2016;**316**:300–312.

Venn A, Watson L, Lumley J, Gilles G, King C, Healy D. Breast and ovarian cancer incidence after infertility and in vitro fertilization. *Lancet* 1995; **346**:995–1000.

Williams CL, Jones ME, Swerdlow AJ, Botting BJ, Davies MC, Jacobs I, Bunch KJ, Murphy MG, Sutcliffe AG. Risks of ovarian, breast and corpus uteri cancer in women treated with assisted reproductive technology in Great Britain, 1991–2010: data linkage study including 2.2 million person years of observation. *Br Med J* 2018;**362**:k2644.