IVF and breast cancer: a systematic review and meta-analysis

Theodoros N. Sergentanis¹, Andreas-Antonios Diamantaras¹, Christina Perlepe¹, Prodromos Kanavidis¹, Alkistis Skalkidou² and Eleni Th. Petridou¹,*

¹Department of Hygiene, Epidemiology and Medical Statistics, Medical School, University of Athens, 75 M. Asias Str. Goudi, Athens 115 27, Greece ²Department of Women’s and Children’s Health, Uppsala University, Uppsala, Sweden

Correspondence address. Tel: +30 210-7462187; Fax: +30 210-7462105; E-mail: epetrid@med.uoa.gr

Submitted on March 30, 2013; resubmitted on June 12, 2013; accepted on June 17, 2013

BACKGROUND: The effects of controlled ovarian hyperstimulation (COH) for IVF in terms of breast cancer risk remain controversial, despite the hormone-dependent nature of the latter.

METHODS: Eligible studies up to 15 February 2013 were identified and pooled effect estimates for relative risk (RR) were calculated separately for the investigations using the general population and those using infertile women, as a reference group. Fixed- or random-effects models were implemented and subgroup analyses were performed, as appropriate.

RESULTS: Eight cohort studies were synthesized, yielding a total cohort size of 1 554 332 women among whom 14 961 incident breast cancer cases occurred, encompassing 576 incident breast cancer cases among women exposed to IVF. No significant association between IVF and breast cancer was observed either in the group of studies treating the general population (RR = 0.91, 95% CI: 0.74 – 1.11) or infertile women (RR = 1.02, 95% CI: 0.88 – 1.18), as a reference group. Of note were the marginal associations, protective for pregnant and/or parous women after IVF (pooled effect estimate = 0.86, 95% CI: 0.73 – 1.01) and adverse for women ≤ 30 years at first IVF treatment (pooled effect estimate = 1.64, 95% CI: 0.96 – 2.80).

CONCLUSIONS: At present, COH for IVF does not seem to impart increased breast cancer risk. Longer follow-up periods, comparisons versus infertile women, subgroup analyses aiming to trace vulnerable subgroups, adjustment for various confounders and larger informative data sets are needed before conclusive statements for the safety of the procedure are reached.

Key words: breast cancer / IVF / infertility / meta-analysis / hormones
**Introduction**

Breast cancer epidemiology has made substantial progress during the recent decades, revealing various risk factors for the disease; their spectrum has spanned hormone-related risk factors, as well as non-hormonal conditions. Hormone-related risk factors have traditionally encompassed endogenous proxies of exposure as well as exogenous exposure to estrogens and progesterone. The inherent factors have included younger age at menarche and older age at menopause (Collaborative Group on Hormonal Factors in Breast Cancer, 2012), post-menopausal obesity (Reenhans et al., 2008; Cheraghi et al., 2012), nulliparity and older age at first birth (Ma et al., 2006), whereas endogenous estradiol levels per se have been associated with increased risk for breast cancer in post-menopausal women (Key et al., 2002) and possibly also in premenopausal women (Eliassen et al., 2006). On the contrary, oophorectomy has been associated with decreased breast cancer risk (Brinton et al., 1988). Recent studies have highlighted the relevance of the inherent hormonal risk factors particularly for tumors positive to hormone receptors, namely estrogen receptors (ER) and progesterone receptors (PR) (Ma et al., 2006; Yang et al., 2011).

The exogenous hormonal factors have mainly encompassed oral contraceptive (OC) use and hormone replacement treatment (HRT) for menopause. Recent meta-analyses pinpointed that HRT is positively associated with increased breast cancer risk (Reeves et al., 2006; Narod, 2011; Marjoniemanks et al., 2012). The Million Women Study has also highlighted that the association involved all histological subtypes, with the largest relative risks (RRs) pertaining to lobular, mixed ductal-lobular and tubular cancers, whereas the effects of HRT were more pronounced for combined estrogen–progesterone therapy than for estrogen-only therapy, as well as among women with normal BMI (Reeves et al., 2006). A multitude of underlying mechanisms has been suggested regarding the potentially aggravating role of the progesterone, such as generation of a population of ER-negative, PR-negative. CKS + cells representing precursors to breast cancer (Horwitz et al., 2008; Horwitz and Sartorius, 2008) and expansion of the mammary stem cell pool (Asselin-Labat et al., 2010; Joshii et al., 2010), whereas the synergistic effects between estrogen and progesterone have been also linked to the estrogen-mediated up-regulation of PR (Cho et al., 1994). Risk of breast cancer seems to increase along with longer HRT usage (Medicines and Healthcare Products Regulatory Agency, 2007); interestingly, according to the Women’s Health Initiative trial, breast cancer risk seems to dissipate within 2 years of cessation of HRT (Chlebowski et al., 2009). Similarly, regarding OC intake, there is evidence that current use may be associated with increased breast cancer risk (Nelson et al., 2012), although such a finding does not seem reproducible upon OC ever use (Nelson et al., 2012; Zhu et al., 2012). Contrary to the above, agents with anti-estrogen properties, such as tamoxifen and raloxifene, correlate with decreased breast cancer risk and seem to be effective in breast cancer prevention (Fisher et al., 2005; Vogel et al., 2010).

On the other hand, previous research has also highlighted the independent significance of inherent or modifiable non-hormonal risk factors for breast cancer, such as older age (Adami et al., 2008), family history (Pharoah et al., 1997) and BRCA1 / BRCA2 mutations (Wang et al., 2012), existence of precursor lesions such as atypical ductal hyperplasia or atypical lobular neoplasia (Marshall et al., 1997) as well as benign breast lesions (Jensen et al., 1989; El-Wakeel and Umpleby, 2003), mammographic density (Nelson et al., 2012; Antoni et al., 2013), lactation (Faupel-Badger et al., 2013), smoking before the first pregnancy (DeRoo et al., 2011; Gaudet et al., 2013), chest radiation during childhood or young adulthood (Travis et al., 2003), in utero exposures (Park et al., 2008), nutritional factors (Taylor et al., 2009; Aune et al., 2012) and genetic variants, such as single nucleotide polymorphisms (Sergentanis and Economopoulos, 2010a, b, c; Zhang et al., 2011). Alcohol consumption may represent a distinct risk factor (Suzuki et al., 2008) integrating hormone-dependent mechanisms, such as modulation of estrogen (Dorgan et al., 2001) and ER levels (Singletary et al., 2001), with hormone-independent actions, such as induction of carcinogenesis by acetaldehyde (Seitz and Stickel, 2007) and reactive oxygen species (Knecht et al., 1990). Similarly, the protective actions mediated by physical activity seem to encompass hormonal and non-hormonal events (Friedenreich and Cust, 2008; Wu et al., 2013). The aforementioned factors may interact with the established hormonal influences; for instance, differential associations of mammographic density (Antoni et al., 2013) with ER and PR subgroups have been observed at the meta-analytical level. Pregnancy-related risk modifiers, such the potentially protective multiple birth (Hsieh et al., 1993; Ji et al., 2007), pre-eclampsia and pregnancy-induced hypertension (Calderon-Margalit et al., 2009a; Nechuta et al., 2010; Opdahl et al., 2012), the rather aggravating high fetal growth of the offspring in terms of weight or length (Wohlfahrt and Melbye, 1999; Xue and Michels, 2007; Cnattingius et al., 2008) and possibly abortions (Brind et al., 1996) may represent a subset of factors integrating hormonal and non-hormonal influences. Interestingly, although pregnancy per se represents a protective factor for breast cancer, a transient increase in breast cancer occurrence after childbirth has been noted, being attributed to the high estrogen and/or progesterone concentrations during pregnancy (Lambe et al., 1994; Liu et al., 2002).

In parallel with the examination of the estrogen/progesterone—breast cancer interplay, the examination of hormonal influences in the context of breast cancer risk gradually expanded to assess the effects of ovarian-stimulating agents for the treatment of infertility. Ovarian-stimulating agents are predominantly used for the treatment of women suffering from World Health Organization ovulation disorders Group I (hypothalamic pituitary failure; namely hypothalamic amenorrhoea or hypogonadotropic hypogonadism) and Group II (hypothalamic pituitary dysfunction; predominately polycystic ovary syndrome) (National Collaborating Centre for Women’s and Children’s Health (UK), 2004; George et al., 2008; Howles et al., 2010); they are administered during the follicular phase of the menstrual cycle so as to increase the serum concentration of gonadotrophins, aiming at follicle maturation and ovulation. Ovarian stimulating agents encompass selective ER modulators, such as clomiphene citrate and tamoxifen (Steiner et al., 2005; Brown et al., 2009; Abu-Hashim, 2012; Misso et al., 2012), urinary or recombinant FSH (Jayaprakasan et al., 2010; Lehert et al., 2010; van Wely et al., 2011; Boostanfar et al., 2012), HCG (Ludwig et al., 2003; Hugues, 2004; Filicori et al., 2005), HMG (Diamond and Wentz, 1986; Daya, 2000; Kelly, 2003) and GnRH analogues (Ron-El et al., 2000; Tay, 2002; National Collaborating Centre for Women’s and Children’s Health (UK), 2004; Messinis, 2005; Humaidan et al., 2012). Ovarian hyperstimulation syndrome (OHSS) (Blankstein et al., 1987) and multiple births (Hack et al., 1970) are considered adverse events of ovulation induction. Nevertheless, significant concerns regarding the potential role of ovarian-stimulating drugs in breast carcinogenesis have arisen (Klip
et al., 2000), given the potentially meaningful modifications in serum estrogen and progesterone levels following the administration of the former (Nakamura et al., 1997; Ayhan et al., 2004). In light of the above, a recent meta-analysis (Zreik et al., 2010) has aimed to evaluate whether ovarian stimulating agents are associated with increased breast cancer risk, quantitatively synthesizing data from 8 case-control and 15 cohort studies. No significant associations were noted either at the overall analysis or the subsets by study design, pharmaceutical agent and combination of drugs or number of cycles.

Although many women have received ovarian-stimulating drugs, usually clomiphene citrate, as a first-line option, the subsequent treatment protocols for COH in the context of IVF are considerably more complex (Polyzos et al., 2010; Youssef et al., 2011a, b; Gibreel et al., 2012). Specifically, gonadotrophins are administered to stimulate ovulation, GnRH analogs aim at shutting down the pituitary to prevent spontaneous ovulation and progesterone is supplemented to counteract GnRH down-regulation which could affect the luteal-phase corpus luteum (Ludwig and Diedrich, 2001; Verberg et al., 2009; Check and Slovis, 2011; Brinton et al., 2013). In terms of breast cancer risk, these combined, potentially synergistic effects may well differ from the isolated effects of ovarian-stimulating agents examined in the studies addressing infertility treatment in general. Moreover, COH in IVF may also entail a distinct temporal aspect, as the elevation in circulating estrogens is only transient; to this end, the literature (Stewart et al., 2012) has critically approached the transient but sizeable increase in estrogen levels, as the latter peak at nearly 3000–4000 pg/ml in an IVF cycle versus only 300 pg/ml in a normal cycle (MacLachlan et al., 1989; Joo et al., 2010).

A host of factors may add to the methodological complexity regarding the assessment of the interaction between COH for IVF and breast cancer risk. Effect estimates describing the association between breast cancer and breast cancer sometimes derive from studies adopting the general population as the reference category, whereas others provide comparisons versus infertile women who were not exposed to IVF; this discrepancy seems to represent a meaningful modifier of the results, as infertility per se is a rather established risk factor for breast cancer (Götin et al., 2008). Aiming to disentangle any tumor-promoting effects, some studies make the distinction between events occurring during the first year of follow-up, thus rendering the synthesis of individual findings even more challenging. Lastly, due to ethical issues, the gold standard well-designed RCTs do not seem applicable to explore this hypothesis; as a result, evidence seems to obligatorily stem from non-randomized study designs.

Given the implications of infertility in terms of public health (Wright et al., 2005; Centers for Disease Control and Prevention, 2010), as well as the upward trends in women seeking and receiving infertility treatment, particularly IVF (Stephen and Chandra, 1998; Nyboe Andersen et al., 2004; Connolly et al., 2009; Kimberly et al., 2012), there is a need for sound and critical synthesis of the existing literature. Methodological issues which may affect the results of this synthesis have been described in detail in our previous effort aiming to explore the role of COH for IVF in relation to ovarian, endometrial and cervical cancer; indeed, synthesis of publications treating general population as the reference group pointed to a statistically significant positive association between IVF and increased risk for ovarian and endometrial cancers, whereas no significant associations with ovarian, endometrial or cervical cancer types were noted when infertile women were used as the reference group (Siriastatidis et al., 2013).

The present systematic review and meta-analysis aims to further pursue and extrapolate our previous effort (Siriastatidis et al., 2013) in the investigation of a potential association between COH for IVF and breast cancer, by disentangling a variety of methodological notions which might modify the effect measures.

**Methods**

The systematic review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Liberati et al., 2009) and in line with the *a priori* protocol agreed upon and signed by all authors.

**Search strategy for the identification of eligible studies**

A broad range search strategy was developed for Ovid Medline (Supplementary data, Fig S1). No restrictions pertaining to language or study design were adopted and the search period spanned from 1966 to 15 February 2013. Reference lists of relevant articles were hand searched for potentially eligible studies (‘snowball’ procedure), so as to maximize the amount of synthesized evidence. Relevant ‘Letters to the Editor’ on previously published or unpublished series were examined for potentially usable data and/or information. The National Institute of Clinical Excellence fertility assessment and treatment guidelines (National Collaborating Centre for Women’s and Children’s Health (UK), 2004) were additionally hand searched. Study authors were contacted for methodological clarifications and provision of missing data. Additionally, corresponding authors of studies that investigated the association between breast cancer and ovulation induction drugs as a whole, or between breast cancer and other assisted reproduction technologies (ARTs), after the year 1981 (the year the widespread use of IVF started) were contacted regarding the availability of IVF-related information, since these studies could possibly have additional unpublished data about IVF procedures. The electronic mail address of the corresponding author, as stated in the articles, was used and in the cases that the delivery of the e-mail failed, alternative e-mail addresses preferably of the corresponding authors or of other coauthors were used. If an author had not responded, a reminder was sent after 1 week and other means of contact, that is telephone and fax, were secondarily used whenever available.

**Study eligibility**

Studies comparing the risk of breast cancer among women undergoing regimens and COH protocols for IVF using the general population or infertile women as reference populations were deemed eligible for this systematic review and meta-analysis. Case series and case reports, *in vitro* and animal studies, narrative or systematic reviews and studies exclusively assessing the treatment of cancer or fertility preservation after cancer treatment were excluded. Studies examining precursor lesions, such as atypical ductal hyperplasia or atypical lobular hyperplasia (lobular neoplasia) were also excluded. Studies about ovarian stimulation for sexual intercourse or intrauterine insemination and not for IVF, as well as studies examining fertility drugs as a whole (when their corresponding authors did not provide us with data especially pertaining to IVF, after being contacted by our team) were not included, given the discrepancies in their protocol from that administered in IVF.

If multiple publications using the same cohort were identified (overlapping studies), the most recent publication was used for data extraction, but information from all relevant publications was retained, if necessary. Two authors (A.A.D. and C.P.) working independently and blindly to each other...
performed the selection of eligible studies, whereas consensus with the whole team was reached in case of disagreement.

**Data extraction**

Two authors (T.N.S. and P.K.) designed and pilot-tested an ad hoc developed excel sheet for data extraction; consensus and approval was subsequently obtained by the whole team of authors. Similarly to our previous meta-analysis (Siristatidis et al., 2013), the abstracted data included general information (title, author, year, journal, geographical and clinical setting), study characteristics (number of participants, design, follow-up, inclusion and exclusion criteria), characteristics of participants (age, ascertainment of exposure and outcome, dose and protocol of IVF, histology, type of infertility, stimulation drugs before IVF, matching factors), details of the statistical analysis (reference population; general population or infertile women adjustment factors at the multivariate analyses, subgroup analyses) and results, i.e. odds ratio (OR), hazard ratio (HR), standardized incidence ratio (SIR), incidence factors at the multivariate analyses, subgroup analyses) and results, i.e. odds ratio (OR), hazard ratio (HR), standardized incidence ratio (SIR), incidence rate ratio (IRR) as reported, and associated raw data for re-calculation (data checking) or de novo estimation of missing measures by our team. Wherever appropriate, maximally adjusted effect estimates were preferred.

Regarding de novo calculations, SIRs were estimated as the ratio of observed over expected number of cases for exposed women. The 95% confidence interval (CI) for log(SIR) was constructed via the term $\pm 1.96/\sqrt{O}$, where $O$ was the observed number of events (Alder et al., 2006; Siristatidis et al., 2013). IRRs and their 95% CIs were estimated from the reported number of incident cases and person-years for exposed and unexposed women, using the epitab STATA commands (StataCorp., 2009).

Two authors (AAD and CP) working independently and blindly to each other performed the data abstraction; again, in case of disagreement, consensus with the whole team was reached.

**Assessment of quality of studies and risk of bias**

The quality of the included studies was evaluated using the nine-item Newcastle-Ottawa Quality scale (Wells et al., 2011). Two authors (TNS and AAD) working independently and blindly to each other performed the assessment of study quality; in case of disagreement, consensus with the whole team was reached. Similarly to the previous effort by our research team, the cut-off level for the adequateness of the follow-up period was set at 10 years (Siristatidis et al., 2013).

Although our initial purpose was to evaluate the existence of publication bias using the Egger’s formal statistical test (Egger et al., 1997), no statistical evaluation was performed given that the number of included studies was small (< 10) and that the power of the test is substantially compromised in this context (Siristatidis et al., 2013).

**Data synthesis**

The effect estimates that were extracted or de novo calculated, namely SIRs, IRRs, HRs and ORs were expected to yield similar estimates of RR given the rarity of the outcome (Larsson et al., 2007; Adami et al., 2008). Consequently, all RR estimates were pooled together (Siristatidis et al., 2013), to ensure the comprehensiveness of the analysis and maximization of the statistical power. Separate analyses were conducted by reference population (general population or infertile women). In addition, subanalyses were presented by type of effect measure, namely SIRs and ORs within the subgroup of studies treating the general population as reference, as well as HRs and IRRs regarding studies treating infertile women as reference group.

Fixed-effects (Mantel–Haenszel) (Mantel and Haenszel, 1959) or random-effects (DerSimonian– Laird) models (DerSimonian and Laird, 1986) were used to calculate pooled effect estimates. Between-study heterogeneity was assessed by using Cochran Q statistic (significance set at 0.1) and by estimating $I^2$ (Higgins et al., 2003; Higgins and Green, 2011). In case of significant heterogeneity, irrespective of the $I^2$ estimation, random effects models were employed (Higgins and Green, 2011). Some of the included studies reported separately data including or excluding incident cases diagnosed during the first year of follow-up; as a result, two distinct approaches were adopted, either preferring effect estimates which excluded the first year of follow-up after IVF or estimates derived from total follow-up (Siristatidis et al., 2013).

In case that more than one study presented relevant subgroup analysis, subgroup analyses were also conducted according to the number of cycles of IVF, histological type of cancer, age group at first exposure to IVF, pregnancy and/or parity, type of subfertility, agent and protocol used for COH, as well as across strata of confounders. Statistical analysis was performed using STATA Software version 11.1 (STATA Corporation, College Station, TX, USA).

**Results**

**Results of the search strategy and contact with corresponding authors**

Figure 1 presents the PRISMA flow chart for the selection of studies. From the 1921 abstracts yielded by the search algorithm, 1861 were excluded based on the title or the abstract; their full-text was necessary for the evaluation of the remaining 60 articles, as well as for another two (Lerner-Geva et al., 2003; Källén et al., 2005a) that were found from the retrieved eligible studies and previous systematic reviews. From these 62 articles only 11 (Venn et al., 1995; Venn et al., 1999; Dor et al., 2002; Lerner-Geva et al., 2003; Källén et al., 2005a; Kristiansson et al., 2007; Pappo et al., 2008; Källén et al., 2011; Stewart et al., 2012; Yli-Kuha et al., 2012; Brinton et al., 2013) studies met the eligibility criteria, 20 articles were investigated further by contacting the authors for additional data (Ron et al., 1987; Brzezinski et al., 1994; Bernstein et al., 1995; Braga et al., 1996; Rossing et al., 1996; Modan et al., 1998; Potashnik et al., 1999; Ricci et al., 1999; Doyle et al., 2002; Petro-Nustas et al., 2002; Burkman et al., 2003; Siegelmann-Danielli et al., 2003; Brinton et al., 2004; Gauthier et al., 2004; Lerner-Geva et al., 2006; Terry et al., 2006; Jensen et al., 2007; Calderon-Margalit et al., 2009b; Orgaas et al., 2009; Silva Idos et al., 2009) and 31 were excluded for various reasons (Talamin et al., 1985; Mauvais-Jarvis, 1987; Rossing and Weiss, 1995; Jourdain et al., 1996; Unkila-Kallio et al., 1997; Duckitt and Templeton, 1998; Rossing and Daling, 1999; Shelley et al., 1999; Cramer et al., 2000; Klip et al., 2000; Venn et al., 2003; Ayhan et al., 2004; Brinton et al., 2005; Salhab et al., 2005; Kanakas and Mantzavinos, 2006; Borini and Rebellato, 2008; Cetin et al., 2008; Katz et al., 2008; Kotsopoulos et al., 2008; Vlahos et al., 2010; Zreik et al., 2010; Bukovic et al., 2011; Brinton, 2012; Fei et al., 2012; Jenks, 2012; Li et al., 2012; Litton, 2012; Turkoz et al., 2012; Twombly, 2012; Li et al., 2013; Russo et al., 2013) (Supplementary data, Table S1).

Regarding the contact with corresponding authors of the 20 studies that examined the association between breast cancer and ovulation induction drugs as a whole, or between breast cancer and other ART after the year 1981, eventually 15 out of 20 authors replied, whereas five (Ron et al., 1987; Rossing et al., 1996; Potashnik et al., 1999; Siegelmann-Danielli et al., 2003; Orgaas et al., 2009) could not be reached or did not reply to any request for additional information; nevertheless, it should be acknowledged that sometimes it is outside...
someone’s control that he/she cannot be reached, given that retirement, death or other events may have occurred.

Despite the prompt reply of 15 authors, further data relevant to the needs of this meta-analysis could not be retrieved, as five authors could not provide information from their study regarding this subject (Braga et al., 1996; Doyle et al., 2002; Burkman et al., 2003; Brinton et al., 2004; Gauthier et al., 2004), three authors (Modan et al., 1998; Lerner-Geva et al., 2006; Calderon-Margalit et al., 2009b) assured that the women participating in their studies had not undergone IVF treatment and six authors (Bernstein et al., 1995; Ricci et al., 1999; Petro-Nustas et al., 2002; Terry et al., 2006; Jensen et al., 2007; Silva Idos et al., 2009) did not examine the history of IVF. Finally, Brzezinski et al. (Brzezinski et al., 1994) disclosed that in their study one breast cancer case was observed in the IVF group (n = 47); however, we opted not to include this study in the meta-analysis, because the expected number of cases could not be calculated with certainty, as it should have been deduced from the general population of the specific region during the study time.

Subsequently, three studies (Venn et al., 1995; Källén et al., 2005a; Kristiansson et al., 2007) were excluded due to overlapping populations with three included studies; specifically, the earliest study by Venn et al. (1995) overlapped with the included, more recent study by the same team (Venn et al., 1999), whereas two studies (Källén et al., 2005a; Kristiansson et al., 2007) overlapped with a more recent included study (Källén et al., 2011). Importantly, after contact with the authors Brinton et al. (2013) reassured that there was no overlapping population between their study and that of Dor et al. (2002), which is the only study performed in Israel that could potentially have overlapped with the former; as a result, both studies were retained in the meta-analysis.

Taken as a whole, eight studies (Venn et al., 1999; Dor et al., 2002; Lerner-Geva et al., 2003; Pappo et al., 2008; Källén et al., 2011; Stewart et al., 2012; Yli-Kuha et al., 2012; Brinton et al., 2013) were...
included in the meta-analysis, representing a total cohort size equal to 1,554,332 women among whom 14,961 incident breast cancer cases were noted. Regarding the potential subanalyses among included studies, Venn et al. (Venn et al., 1999) provided additional subgroup analyses by type of infertility, while Stewart et al. (Stewart et al., 2012) kindly ensured that age at start of infertility investigation did not coincide with age at first actual infertility treatment for the total sample, but could not provide estimates which excluded the first year of follow-up after IVF (Venn et al., 2012; Stewart et al., 2013). Furthermore, Källén et al. kindly supplied us with the risk ratio pertaining to women exposed to IVF before 30 years of age (Källén et al., 2011).

Table I presents the details of the eight included cohort studies; four of them were conducted in Israel (Dor et al., 2002; Lerner-Geva et al., 2003; Pappo et al., 2008; Brinton et al., 2013), two in Australia (Venn et al., 1999; Stewart et al., 2012), one in Finland (Yli-Kuha et al., 2012) and one in Sweden (Källén et al., 2011), encompassing 576 incident breast cancer cases among women exposed to IVF. Five studies presented comparisons versus the general population (Dor et al., 2002; Lerner-Geva et al., 2003; Pappo et al., 2008; Källén et al., 2011; Yli-Kuha et al., 2012), two studies versus infertile women (Stewart et al., 2012; Brinton et al., 2013), whereas in one study both comparisons were applied (Venn et al., 1999).

Interestingly, case admixture with in situ lesions was not uniformly treated by the individual studies; specifically, the majority of included studies did not disclose the percentage of in situ lesions (Dor et al., 2002; Lerner-Geva et al., 2003; Källén et al., 2011; Yli-Kuha et al., 2012; Brinton et al., 2013) whereas in the remaining studies the percentage ranged between 2.1% (Venn et al., 1999) and 14.3% (Stewart et al., 2012).

Quality of the included studies
Rating of the quality of studies according to the Newcastle-Ottawa score is presented in Supplementary data, Table SII, while the PRISMA Checklist in Supplementary data, Table SIII. The quality scores ranged between 7 and 9. The main shortcoming of the included studies pertained to the rather short follow-up period, as only one study (Stewart et al., 2012) was based on a follow-up period >10 years including exposed women. Regarding the comparability of groups, all studies ensured the comparability on age, whereas four studies provided the comparability on additional factors. Specifically, Källén et al. additionally adjusted for year of delivery and smoking (Källén et al., 2011), Stewart et al. additionally adjusted for age at first birth as well as delivery of twins and higher-order multiples (Stewart et al., 2012), Yli-Kuha et al. provided matching on residence as well as adjustment for marital status and socioeconomic position (Yli-Kuha et al., 2012) and Brinton et al. also adjusted for BMI, smoking, parity and socioeconomic status, including also a term in the model for the fertility treatment other than IVF (GnRH analogues, clomiphene, progestogen) (Brinton et al., 2013).

Synthesis of studies
The synthesis of studies using two alternative approaches (preferring estimates which excluded the first year of follow-up after IVF, as well as those derived from total follow-up) are presented in Table II. At the approach preferring estimates which excluded the first year of follow-up after IVF (upper panels), studies treating the general population as the reference category did not point to a statistically significant association between IVF and breast cancer risk (pooled effect estimate = 0.91, 95% CI: 0.74–1.11, Fig. 2a). Similarly, the synthesis of studies treating infertile women as the reference group did not suggest a significant correlation (pooled effect estimate = 1.02, 95% CI: 0.88–1.18, Fig. 2b).

Subanalyses by the type of effect estimate that studies provided resulted in the emergence of an inverse association between IVF and breast cancer risk in studies versus the general population which provided ORs (pooled effect estimate = 0.78, 95% CI: 0.65–0.94), but not in studies providing SIRs (pooled effect estimate = 0.99, 95% CI: 0.73–1.34); the respective subanalyses are illustrated in Fig. 3. The subanalyses within the subgroup of studies treating infertile women as the reference group are presented in Supplementary data, Fig. S2; no significant associations were noted therein.

The alternative approach preferring estimates derived from the total follow-up essentially replicated the results of the aforementioned analyses based on estimates which preferred those excluding the first year of follow-up after IVF. The respective forest plots are presented in Supplementary data, Figs S3–S6.

Supplementary data, Table SIV illustrates the subanalyses that each study contributed; notably, the table contains only five studies (Venn et al., 1999; Pappo et al., 2008; Källén et al., 2011; Stewart et al., 2012; Brinton et al., 2013), as the remaining ones did not present any subgroup analyses. The three subclassifications that were separately examined by more than one study pertained to number of IVF cycles, age at first IVF treatment as well as pregnancy and/or parity after IVF. Specifically, three studies presented subgroup analyses on number of IVF cycles (Venn et al., 1999; Pappo et al., 2008; Brinton et al., 2013); the synthesis of the effect estimates comparing the highest exposure category (largest number of IVF cycles) versus not exposed women did not point to increased risk for breast cancer after IVF (pooled effect estimate = 1.28, 95% CI: 0.77–2.15, Fig. 4a). Regarding pregnancy and/or parity (Pappo et al., 2008; Källén et al., 2011; Brinton et al., 2013), no significant association was observed among nulliparous women (pooled effect estimate = 1.04, 95% CI: 0.51–2.13, Supplementary data, Figure S7), whereas a marginally protective trend among women pregnant and/or parous after IVF (pooled effect estimate = 0.86, 95% CI: 0.73–1.01, Fig. 4b) did not reach formal significance. Similarly, no excess risk was noted among women aged ≥30 years at first IVF treatment (pooled effect estimate = 1.18, 95% CI: 0.72–1.94, Fig. 4c); the positive association regarding women <30 years at first IVF treatment did not reach significance (pooled effect estimate = 1.64, 95% CI: 0.96–2.80, Supplementary data, Fig. S8).

Discussion
The meta-analysis of currently available studies pointed to an overall lack of association between COH for IVF and subsequent breast cancer risk. The null association was reproducible upon the alternatively performed analyses, namely in both approaches regarding exclusion or inclusion of early events during the follow-up period, as well as studies treating as reference the general population or infertile women. Synthesis of the few study arms examining women exposed to the largest number of IVF cycles in order to explore dose–response effects underlying causation as per the Bradford-Hill criteria (Hill, 1965) seemed also to extrapolate
Table 1 Characteristics of included studies in the systematic review and meta-analysis of the impact of IVF treatment on breast cancer.

<table>
<thead>
<tr>
<th>Study publication</th>
<th>Country, region</th>
<th>Study period (including follow-up)</th>
<th>Cohort size</th>
<th>Total number of exposed women</th>
<th>Number of incident cases</th>
<th>Number of exposed cases</th>
<th>Mean follow-up in total cohort (years)</th>
<th>Mean follow-up in exposed women (years)</th>
<th>Study protocol for IVF</th>
<th>Effect estimates</th>
<th>Reference group</th>
<th>Adjusting factors</th>
<th>Excludes first year of follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dor et al. (2002)</td>
<td>Israel (Tel Hashomer, Tel Aviv)</td>
<td>1981–1996</td>
<td>5026</td>
<td>5026</td>
<td>11</td>
<td>11</td>
<td>3.6 ± 3.4</td>
<td>3.6 ± 3.4</td>
<td>1. CC/hMG, FSH, SIR LH 2. hMG 3. GnRH-agonist/ hMG</td>
<td>General population</td>
<td>None</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Källén et al. (2011)</td>
<td>Sweden (all IVF clinics)</td>
<td>1982–2006</td>
<td>1 388 371</td>
<td>23 192</td>
<td>13 746</td>
<td>95</td>
<td>NR</td>
<td>8.3</td>
<td>NR</td>
<td>OR</td>
<td>General population</td>
<td>Year of delivery, maternal age at delivery and smoking</td>
<td>No</td>
</tr>
<tr>
<td>Lerner-Geva et al. (2003)</td>
<td>Israel (Tel Aviv)</td>
<td>1984–1996</td>
<td>1082</td>
<td>1082</td>
<td>5</td>
<td>5</td>
<td>6.5 ± 2.2</td>
<td>6.5 ± 2.2</td>
<td>NR</td>
<td>SIR</td>
<td>General population</td>
<td>None</td>
<td>Both</td>
</tr>
<tr>
<td>Yli-Kuha et al. (2012)</td>
<td>Finland</td>
<td>1996–2004</td>
<td>18 350</td>
<td>9175</td>
<td>115</td>
<td>55</td>
<td>7.8</td>
<td>7.8</td>
<td>NR</td>
<td>OR</td>
<td>General population</td>
<td>Socioeconomic position and marital status</td>
<td>Both</td>
</tr>
<tr>
<td>Pappo et al. (2008 )</td>
<td>Israel (Tel Aviv, 1 clinic)</td>
<td>1986–2004</td>
<td>3375</td>
<td>3375</td>
<td>35</td>
<td>35</td>
<td>8.1 ± 4.3</td>
<td>8.1 ± 4.3</td>
<td>NR</td>
<td>SIR</td>
<td>General population</td>
<td>Age, continent of birth</td>
<td>No</td>
</tr>
<tr>
<td>Stewart et al. (2012)</td>
<td>Western Australia (all clinics)</td>
<td>1983–2010</td>
<td>21 025</td>
<td>7381</td>
<td>384</td>
<td>148</td>
<td>16.3 ± 5.6</td>
<td>16.1 ± 5.6</td>
<td>NR</td>
<td>HR</td>
<td>Infertile women</td>
<td>Age at first birth and the delivery of twins and higher order multiples</td>
<td>Yes (first 6 months)</td>
</tr>
<tr>
<td>Brinton et al. (2013)</td>
<td>Israel (25% of total population)</td>
<td>1994–2011</td>
<td>87 403</td>
<td>NR</td>
<td>522</td>
<td>140</td>
<td>8.1</td>
<td>8.1</td>
<td>NR</td>
<td>HR</td>
<td>Infertile women</td>
<td>Age at entry, BMI, Both* smoking, parity at exit and socioeconomic status</td>
<td></td>
</tr>
</tbody>
</table>

Study publication: Mean age in total cohort, Mean age in exposed cohort, Cohort characteristics, Ascertainment of exposure, Ascertainment of cancer, Histology, In situ lesions % (n), Type of infertility.

Dor et al. (2002): 34.0 ± 6.4 at first treatment; 37.5 ± 7.1 at end of follow-up; Exposed: treated for subfertility and had at least 1 cycle of IVF; Medical records; Israel National Cancer Registry; NR; NR; Data only for the First Department (1254 women overall): 48.7% mechanical, 8.6% ovulatory, 19.4% male factor, 23.3% unexplained.

Källén et al. (2011): 32.0 at first delivery; 40.3 at end of follow-up; Exposed: women who delivered an infant following IVF treatment; Medical records; National Board of Health and Welfare; Swedish Cancer Registry; NR; NR; NR.

Lerner-Geva et al. (2003): 32.7 ± 4.8 at first treatment; 38.7 ± 5.2 at end of follow-up; Exposed: diagnosed with subfertility problems and had at least 1 cycle of IVF; Medical records; Israel National Cancer Registry; NR; NR; 42.2% mechanical, 24.2% hormonal, 30.1% male, 3.5% unexplained.
<table>
<thead>
<tr>
<th>Study</th>
<th>Median Age at Entry</th>
<th>Median Age at End of Follow-up</th>
<th>Exposure Details</th>
<th>Medical Records/Database Details</th>
<th>Cancer Types</th>
<th>Treatment Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Venn et al. (1999)</td>
<td>30.7 at entry</td>
<td>39.9 at end of follow-up</td>
<td>Evaluated for subfertility and had at least one IVF treatment cycle with ovarian stimulation (including stimulated cycles that were cancelled). Unexposed: referred for IVF but untreated or had 'natural cycle' treatment without ovarian stimulation.</td>
<td>Medical records, computerized data for four clinics. State population-based Cancer Registries, National Cancer Statistics Clearing House and National Death Index.</td>
<td>71% invasive ductal carcinomas</td>
<td>2.1% (3) DCIS, 40.0% tubal, 33.9% male factor, 17.3% endometriosis, 5.2% ovarian defect, 4.4% other, 12.6% unexplained, 11.4% missing.</td>
</tr>
<tr>
<td>Ylikuha et al. (2012)</td>
<td>33.5 at first treatment</td>
<td>33.5 at first treatment</td>
<td>Received IVF (including ICSI and FET); Unexposed: population register (matched by age, municipality).</td>
<td>Identified by reimbursement for drugs or drug combinations specific to these treatments. Finnish Cancer Registry.</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Pappo et al. (2008)</td>
<td>32.1 ± 5.7 at first treatment</td>
<td>32.1 ± 5.7 at first treatment</td>
<td>Treated for subfertility and had at least 1 cycle of IVF.</td>
<td>Medical records. Israel National Cancer Registry.</td>
<td>66% IDC, 9% ILC, 14% DCIS, 5.5% tubular, 5.5% mucinous</td>
<td>10.5% hormonal, 86.1% non-hormonal.</td>
</tr>
<tr>
<td>Stewart et al. (2012)</td>
<td>31.3 ± 5.1 at first fertility investigation, 48.0 ± 6.9 at end of follow-up</td>
<td>32.1 ± 4.8 at first fertility investigation, 48.6 ± 6.7 at end of follow-up</td>
<td>Women at their first infertility admission (for IVF treatment or not) with diagnosis of either infertility or procreative management.</td>
<td>Women at their first infertility admission (for IVF treatment or not) with diagnosis of either infertility or procreative management. Reproductive Technology Register, Medical Records.</td>
<td>WA Cancer Registry</td>
<td>14.3% (55)</td>
</tr>
<tr>
<td>Brinton et al. (2013)</td>
<td>31.1 at entry</td>
<td>NR</td>
<td>Exposed: treated for infertility and had at least 1 IVF cycle: unexposed: registered with fertility problems and did not have infertility treatment.</td>
<td>Maccabi Healthcare Services. Israel National Cancer Registry.</td>
<td>NR</td>
<td>21.4% male infertility, 15.3% anovulation, 6.6% mechanical, 2.8% PCOS, 0.3% pituitary-hypothalamic, 1.9% endometriosis, 50.9% unspecified.</td>
</tr>
</tbody>
</table>

HR, hazard ratio; IRR, incidence rate ratio; NR, not reported; OR, odds ratio; SIR, standardized incidence ratio; CC, clomiphene citrate; FET, frozen embryo transfer; DCIS, ductal carcinoma in situ; LCIS, lobular carcinoma in situ; IDC, invasive ductal carcinoma; ILC, invasive lobular carcinoma; PCOS, polycystic ovary syndrome; ICSI, intracytoplasmic sperm injection.

*Brinton et al. kindly provided the additional results excluding breast cancer cases occurring during the first year of follow-up.
the null association upon this theoretically most vulnerable subpopulation. Worth noting was a marginally protective effect among women pregnant and/or parous after IVF; a sizeable adverse association underlines the challenges that emerge at the approach to the association outside the context of IVF. In the present meta-analysis, both subsets addressing comparisons versus infertile women, only one provided an adjustment term in the model for the fertility treatment (Brinton et al., 2013) and another one solely acknowledged this shortcoming as a limitation (Stewart et al., 2012). Consequently, the optimal study design would theoretically envisage comparisons versus infertile women adjusting or controlling for exposure to other ovulation induction agents outside the context of IVF. In the present meta-analysis, both subsets of studies, irrespectively of the reference category, pointed to a null effect of IVF in terms of breast cancer risk.

Another methodological point in the synthesis and interpretation of the study findings pertains to the importance of early events during the follow-up period, especially given the established temporarily increased risk for breast cancer following pregnancy (Lambe et al., 1994; Liu et al., 2002) and the consequent difficulty to attribute it to the pregnancy itself or the IVF treatment. Moreover, early events may suggest tumor-promoting effects rather than true causation, whereas diagnostic access bias may be associated with higher diagnosis rates during the
close contact of women receiving IVF with the health services. Both aforementioned axes would theoretically lead to the overestimation of the effect estimates; as a result, four out of eight eligible studies made this interesting distinction (Dor et al., 2002; Lerner-Geva et al., 2003; Yli-Kuha et al., 2012; Brinton et al., 2013). In our synthesis, such modifying results did not seem to play a major role, possibly due to the rarity of the events; indeed, the effect estimates derived from both approaches regarding the exclusion or inclusion of events occurring during the first follow-up year replicated each other. This may be considered a hallmark of the internal consistency of results, although ideally all included studies should have presented such subanalyses, so as to maximize the statistical power of the alternative approaches.

Regarding the variability of effect estimates in this meta-analysis, pooling of SIRs and ORs, as well as HRs and IRRs was performed in the synthesis of studies versus the general population and infertile women, respectively. Pooling was allowed from a statistical point of view, given their asymptotic convergence to RR assuming the relative rarity of the outcome variable (Larsson et al., 2007; Adami et al., 2008; Siristatidis et al., 2013). Nevertheless, subgroup analyses by effect estimate were always presented in order to ensure detailed reporting and examination of potential differences along the effect estimates, but the results should be interpreted with caution, as the number of studies therein was small.

In view of the paucity of primary data and the complexities of the related studies, the aforementioned methodological points should ideally be taken into account in future meta-analyses aiming at accurate assessment of a tentative cancer risk imparted by IVF in order to avoid misinterpretations in summarizing published literature. Indeed, a recently published meta-analysis aiming to examine the effects mediated by IVF regarding cancer risk, with emphasis on ovarian, breast and cervical cancer (Li et al., 2013) seems to have been hampered by methodological limitations (Sergentanis et al., under review), given that two pairs of mutually overlapping studies—namely one pair in Sweden (Kristiansson et al., 2007; Källén et al., 2011) and one pair in Australia (Venn et al., 1995; Venn et al., 1999)—had erroneously been retained, representing essentially the duplication of information. Moreover, the aforementioned meta-analysis has not addressed the subtle but potentially decisive implications of reference category, early events and effect estimates.

Enlarging the perspective, the results of the present meta-analysis seem in accordance with those of our previous meta-analysis, according
to which COH for IVF was not associated with increased risk for endometrial, ovarian or cervical cancer in the most meaningful comparisons versus infertile women. Women after IVF may run the risk of short-term conditions, such as OHSS (Delvigne and Rozenberg, 2002; Chan and Dixon, 2008; D’Angelo et al., 2011; Venetis et al., 2011), but both our recent meta-analyses suggest that the concerns for elevated cancer risk after IVF do not seem particularly justified, based on the current evidence. On the other side of the mother–child dipole; however, conception of children with ART has been associated with increased risk of multiple pregnancy (Basatemur and Sutcliffe, 2008; Nyboe Andersen et al., 2009) especially prior to the advent of single embryo transfer (Gelbaya et al., 2010; McLernon et al., 2010), major birth defects and congenital malformations (Kurinczuk et al., 2004; Bonduelle et al., 2005; Klemetti et al., 2005; Lie et al., 2005; Olson et al., 2005; Rimm et al., 2011; Hansen et al., 2012), preterm birth (Helmerhorst et al., 2004; Blickstein, 2006; Henningsson et al., 2011; Sazonova et al., 2011; Grady et al., 2012), cerebral palsy (Hvidtjørn et al., 2009; Källén et al., 2010a; Zhu et al., 2010), low birthweight (McDonald et al., 2009, 2010), whereas concerns have been raised regarding genetically inherited syndromes (Ludwig et al., 2005; Sutcliffe et al., 2006; Lim et al., 2009), endocrine and metabolic conditions (Ceelen et al., 2007; Ceelen et al., 2008; Belva et al., 2012) including subclinical hypothyroidism (Sakka et al., 2009; Onal et al., 2012) or even childhood cancer (Moll et al., 2003; Källén et al., 2005b; McLaughlin et al., 2006; Källén et al., 2010b; Akefeldt et al., 2012; Petridou et al., 2012; Rudant et al., 2013). Therefore, the health effects mediated by IVF on women and the offspring seem to present an intriguing discrepancy and underlying mechanisms should be explored.

The limitations of this meta-analysis necessarily reflect the inherent limitations of the constituent studies, as reflected in their quality ratings. The major aspect interfering with the quality of the included studies pertained to the rather short follow-up period, as only one study (Stewart et al., 2012) encompassed a follow-up period > 10 years among exposed women; given that exposure to IVF mainly occurs during the late reproductive years, an adequate follow-up period seems indispensable to effectively assess particularly the risk of post-menopausal breast cancer. Furthermore, many studies presented SIRs which inherently correspond to RR estimates adjusted only for

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure3}
\caption{Subanalyses within studies treating the general population as a reference group, preferring estimates excluding the first year of follow-up after IVF. Forest plots presenting combined effect estimates for breast cancer in women exposed to IVF, in studies (a) presenting SIRs and (b) presenting ORs. (SIR, standardized incidence ratio; OR, odds ratio.)}
\end{figure}
age and calendar time, whereas IRRs are based on crude estimates; as a result, only four studies controlled for factors additional to age (Källén et al., 2011; Stewart et al., 2012; Yli-Kuha et al., 2012; Brinton et al., 2013). The potential effect of residual confounding by other factors which have not been controlled for, such as differential patterns of breast feeding among IVF women (Hammarberg et al., 2011), should also be kept in mind.

Regarding the actual size of the data set for this meta-analysis, despite the meticulous evaluation of nearly 2000 abstracts, only 8 cohort studies were eligible; case–control studies did not provide data pertaining

Figure 4  Forest plots presenting subgroup analyses for the association between IVF and breast cancer risk. (a) Number of IVF cycles; synthesis of the effect estimates comparing the highest exposure category versus not exposed women; (b) subgroup analysis on women pregnant and/or parous after IVF and (c) subgroup analysis on study arms including women aged ≥30 years at first IVF treatment. (ES, effect size.)
especially to IVF, as they examined the overall association between fertility drugs and breast cancer. Among the cohort studies, solely five (Venn et al., 1999; Pappo et al., 2008; Källén et al., 2011; Stewart et al., 2012; Brinton et al., 2013) presented subgroup analyses. Consequently, the statistical power of subgroup analyses pertaining to IVF cycles, pregnancy and/or parity after IVF and age at first IVF treatment may have been hampered owing to the fact that only few studies presented the respective subgroup analyses. Indeed, interactions between subgroups, such as effects among women >40 years of age receiving four or more IVF cycles, that have been described as rather devastating with 8.6-fold increased breast cancer risk (Pappo et al., 2008), could not be assessed at the meta-analytical level due to the paucity of data. Moreover, meaningful subgroup analyses, such as those pertaining to the fertility drugs used, could not be addressed, since only one study provided the relevant data (Venn et al., 1999).

The effect estimate describing the association between IVF and breast cancer risk among women <30 years at first IVF treatment, based only on two studies (Pappo et al., 2008; Källén et al., 2011), was relatively sizeable (pooled RR = 1.64) but did not reach statistical significance. This finding seems in line with the observation by Stewart et al. (2012) who supported excess risk among women commencing infertility investigation before 24 years of age, but it should be kept in mind that in our meta-analysis one of the two synthesized effect estimates (Källén et al., 2011) was an unadjusted risk ratio obtained after contact with the corresponding authors. The marginally protective effect among parous women should be also interpreted with caution, on account of the need for further accumulation of evidence as well as stratification by reference group and type of effect estimate; those procedures could not be performed in our setting, given the paucity of study arms.

In addition, case admixture with in situ lesions was not disclosed by the majority of individual studies; this hampered the potential of this meta-analysis to examine overdiagnosis as a proxy of diagnostic access bias. Furthermore, separate analyses were not provided by histological subtypes of breast cancer or by ER and PR status, the latter being potentially of special interest, as discrepancies have been noted regarding the actions of traditional hormonal risk factors by hormone receptor status (Ma et al., 2006; Yang et al., 2011).

On the other hand however, among the strengths and assets of this meta-analysis, numerous aspects may be counted, in addition to the elaborate methodological approach. The rigorous contact with the authors of the individual studies, a step often missed by systematic reviewers (Mullan et al., 2009), has explored and reached the limits regarding the pursuit of potentially existing, additional evidence. Clear definitions of exposure and outcomes, no language restriction and adherence to procedures that minimize extraction, recording and retrieval bias may also be included in the advantages of this study.

In conclusion, the synthesis of currently available data suggests that IVF is not associated with increased breast cancer risk, either when the reference group is the general population or when the comparison is made versus infertile women; exclusion of early cancer events during follow-up does not seem to affect results. Future cohort studies should ideally span longer follow-up periods, provide comparisons versus infertile women and encompass adjustment for meaningful confounders among which exposure to ovulation induction outside the context of IVF seems pivotal. Detailed results concerning subgroup analyses in an attempt to trace vulnerable subgroups are also particularly useful. Larger informative data sets are needed before conclusive statements for the safety of the procedure are reached with confidence.

**Supplementary data**

Supplementary data are available at http://humupd.oxfordjournals.org/.

**Acknowledgements**

We would like to thank the corresponding authors of 15 studies who replied to our request for additional data, as detailed in the manuscript. We would also like to thank the authors of included studies Dr Brinton, Dr Källén, Dr Stewart and Dr Venn for the provision of additional information. This work was partially supported by the National and Kapodistrian University of Athens.

**Authors’ roles**

T.N.S. conceived the idea of the study, contributed to study design, critical evaluation of the studies, extraction and interpretation of data, performed statistical analysis, drafted the article and gave final approval of the version to be published; A.A.D. contributed to study design, critical evaluation of the studies, extraction and interpretation of data, performed statistical analysis, drafted the article and gave final approval of the version to be published; C.P. contributed to study design, critical evaluation of the studies, extraction and interpretation of data, drafted the article and gave final approval of the version to be published; P.K. contributed to study design, critical evaluation of the studies, extraction and interpretation of data, performed statistical analysis, drafted the article and gave final approval of the version to be published; A.S. contributed to study design, critical evaluation of the studies, interpretation of data, drafted the article and gave final approval of the version to be published; E.T.P. conceived the idea of the study, designed the study, contributed to critical evaluation of the studies, extraction and interpretation of data, performed statistical analysis, drafted the article and secured the invitation, gave final approval of the version to be published and will act as a guarantor of the study.

**Funding**

All authors declare no other relationship with funding sources or sponsorships.

**Conflict of interest**

All authors declare no conflict of interest.

**References**


