Effects of assisted reproductive techniques on offspring gonadal function: a systematic review and meta-analysis

Andrea Crafa, M.D., Rossella Cannarella, M.D., Ph.D., Federica Barbagallo, M.D., Sandro La Vignera, M.D., Ph.D., Rosita A. Condorelli, M.D., Ph.D., and Aldo E. Calogero, M.D.

Department of Clinical and Experimental Medicine, University of Catania, Catania, Italy.

Objective: To evaluate whether the use of assisted reproductive techniques (ARTs) affected the gonadal function of the offspring. Numerous concerns have emerged over the years about the use of ARTs and their effects on the health of the offspring.

Evidence Review: Data were extracted through extensive searches in the PubMed and Scopus databases from their establishment until August 2022. Meta-analysis was performed following the Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols guidelines. All eligible studies were selected according to the Population, Exposure, Comparison/Comparator, Outcomes, Study design model. All studies that analyzed pubertal development and testicular and ovarian function of offspring conceived using ARTs were included. The quality of the studies was assessed using the Cambridge Quality Checklists. For the different outcomes, the standardized mean difference (SMD), the mean difference, and the odds ratio were evaluated. Cochran-Q and I^2 statistics were used to evaluate statistical heterogeneity. Differences in pubertal development, hormone levels, and sperm function in ART-conceived subjects compared with spontaneously conceived (SC) control subjects.

Results: Children conceived using ART do not appear to have impaired pubertal development or achievement of pubertal milestones. From an endocrine point of view, ART-conceived males showed lower sex hormone-binding globulin levels than the control group (SMD = −0.25 [−0.44, −0.05]; I^2 = 29%) and a tendency to lower testosterone levels (SMD = −0.16 [−0.32, 0.01]; I^2 = 0). Lower levels of inhibin B, on the other hand, were present only in intracytoplasmic sperm injection-born males compared with the control group (SMD = −0.24 [−0.44, −0.04]; I^2 = 0). In females, higher luteinizing hormone levels were found in the ART group than in the control group (SMD = 0.33 [0.06, 0.59]; I^2 = 17%). In contrast, lower levels of 17ß-estradiol were observed only in the intracytoplasmic sperm injection group compared with girls with SC (SMD = −0.39 [−0.74, −0.03]; I^2 = 0). However, no definitive conclusions can be drawn considering the heterogeneity of hormonal assessments in females. Finally, young adults born from ART had a reduced sperm concentration (SMD = −0.34 [−0.57, −0.11]; I^2 = 0), total sperm count (SMD = −0.28 [−0.51, −0.05]; I^2 = 0), and normal sperm morphology (SMD = −0.35 [−0.58, −0.13]; I^2 = 0) compared with those SC.

Conclusion: A slight alteration in the function of the male germinal epithelium appears to be associated with the use of ART, as shown by the reduced levels of inhibin B and the altered sperm parameters. (Fertil Steril Rev® 2023;4:152–73. ©2023 by American Society for Reproductive Medicine.)

Key Words: Assisted reproductive technique, offspring, pubertal development, gonadal function, sperm parameters

ESSENTIAL POINTS

- To our knowledge, this is the first systematic review and meta-analysis attempting to evaluate the effects of assisted reproductive technique (ART) on offspring gonadal function.
- Leydig cell function appears to be preserved in the offspring conceived using ART.
- The spermatogenic function appears to be worse in offspring conceived using ART.

Received November 1, 2022; revised March 7, 2023; accepted March 8, 2023.

A.C. has nothing to disclose. R.C. has nothing to disclose. F.B. has nothing to disclose. S.L.V. has nothing to disclose. R.A.C. has nothing to disclose. A.E.C. has nothing to disclose.

Correspondence: Rossella Cannarella, M.D., Ph.D., Department of Clinical and Experimental Medicine, University of Catania, via S. Sofia 78, Catania, Italy 95123 (E-mail: rossella.cannarella@phd.unict.it).

Fertil Steril Rev® Vol. 4, No. 2, April 2023 2666-5719/$36.00

https://doi.org/10.1016/j.xfnr.2023.03.001

Infertility is a condition that affects approximately 15% of couples in the world, with a significant economic impact on the health care system largely related to the use of assisted reproductive techniques (ARTs) (1). In fact, because of Louise Brown's birth with conventional in vitro fertilization (IVF) in 1978, the use of ART has increased gradually to the point that >5 million infants have been conceived with these techniques. In some countries, the rate of infants born using ART is >5% (2).

However, numerous concerns have emerged over the years about the use of these techniques and their effect on the health of offspring. According to Barker's (3) theory of the developmental origins of health and disease, prenatal exposure to an altered microenvironment could be associated with long-term disease. Indeed, an association of these techniques with adverse obstetric and perinatal outcomes has been suggested, such as the risk of low birth weight, preterm delivery, small-for-gestational-age (SGA), stillbirth, perinatal mortality, and admission to a neonatal intensive care unit (4). It has also been suggested that the microenvironment plays a role in increasing the risk of birth defects, cancers, growth and development disorders, and chronic diseases, such as diabetes mellitus, obesity, and cardiovascular disease (4).

In particular, the interference of the procedures used for ART on DNA methylation reprogramming during gametogenesis and embryogenesis has been proposed as the underlying mechanism for these alterations. A recent meta-analysis examined this issue, concluding that the methylation pattern of offspring conceived after ART may be different from that of spontaneously conceived (SC) offspring. However, the heterogeneity of the studies and the lack of prospective studies of adequate size did not allow definitive conclusions to be drawn in this regard (5). To confirm the epigenome's possible alteration, meta-analytic studies have shown an increased prevalence of imprinting disorders, such as Beckwith-Wiedemann, Angelman, Prader-Willi, and Silver-Russell syndromes in newborns conceived using ART (6, 7).

Although several studies have evaluated the impact of ART on the metabolic, cardiovascular, and psychoneurologic development of offspring (4), there is little evidence on the impact of these techniques on pubertal development and reproductive function of the offspring. However, it has been suggested that these functions may also be impaired. Indeed, a recent meta-analysis showed an increase in the rate of urogenital abnormalities (cryptorchidism and hypospadias), in newborns conceived using ARTs. These conditions, in turn, may be associated with abnormalities in reproductive function (8). It has been also observed that the expression of genes involved in reproductive function, such as spermatogenesis and centriole-associated 1–like gene that encodes for speriolin, could also be altered in ART-conceived offspring (9). This protein is involved in the formation of the sperm centriole. An alteration of the methylation of the gene that encodes it could, therefore, alter the formation of the centriole, which in turn plays a fundamental role in the correct motility of spermatozoa and the formation of the first centrosome of the embryo (10). Furthermore, alterations in the formation and quality of this centriole have been associated with idiopathic infertility (10). Finally, a recent cohort study involving a total of 122,321 children conceived using ART and 6,576,410 children conceived naturally in Denmark, Finland, Norway, and Sweden found a higher rate of diagnoses related to puberty disorders in the former than in the latter. In particular, children conceived using ART appear to be at a higher risk of precocious and delayed puberty even after adjusting for various risk factors, including gestational age, which is one of the main mechanisms proposed for the alteration of puberty in these children. This suggests a direct role of ART, even if the effects mediated by underlying parental infertility cannot be excluded (11). Taking these considerations into account, the reproductive function of offspring conceived using ART is undoubtedly an aspect that deserves further and more in-depth evaluation.

This systematic review and meta-analysis were undertaken to examine all the evidence present to date in the literature regarding the association between ART and pubertal development and hormonal and reproductive function.

MATERIALS AND METHODS

This study was conducted using the preferred reporting items for systematic review and meta-analysis protocols guidelines (12) (Supplemental Table 1, available online).

Search Strategy

The data were extracted through extensive searches of the PubMed, Scopus, Web of Science, and Cochrane reviews databases from their creation through December 2022. The search strategy included a combination of several medical subjects headings terms and keywords. In particular, the following medical subjects headings terms were used: "Reproductive Techniques, Assisted," "Sperm Injections, Intracytoplasmic," "Fertilization in Vitro," "Puberty," and "Hormone." Moreover, the following keywords were also used: "in vitro fertilization," "IVF," "intrauterine insemination," "IUI," "offspring," "follicular count," "sperm," and "ICSI." For other keywords, whenever possible, the asterisk operator has been used to include several terms simultaneously. In detail, the keyword "assisted reproductive techn*" was used to include the terms assisted reproductive technique, assisted reproductive technology, and assisted reproductive technics. Similarly, using the search terms "pubert*" and "hormon*," we wanted to include the terms puberty, pubertal, hormone, and hormonal at the same time. The specific search strings used for each database queried are reported in Supplemental Table 2 (available online).

Initially, studies were evaluated for inclusion by reading their abstracts. Only articles in English were selected.

When the abstract was not sufficient to determine whether the article-contained data are useful for the meta-analysis, the full texts of the articles were evaluated carefully. The identification of eligible studies was performed independently by 2 different researchers (A.C. and R.C.). Any disagreements were resolved by a third author (A.E.C.). Other articles were extracted manually by searching the reference lists of the selected articles using the keywords indicated above.

Inclusion and Exclusion Criteria

All eligible studies were selected according to the Population, Exposure, Comparison/Comparator, Outcomes, Study design model (Table 1). All studies reporting the effects of ART on pubertal development or hormonal or reproductive function of the offspring were included. Case reports, comments, letters to the editor, systematic or narrative reviews, animal studies, and those studies that did not allow to extract the outcomes of interest were excluded from the analysis. Studies in which the offspring were born from infertile parents who did not undergo infertility treatment were also excluded.

Data Extraction

Data extraction was performed by one author (A.C.) and verified by a second one (R.C.). The following data were collected: the first author's name, age of publication, study design, the total number of patients (including the respective controls) for each outcome, the age of patients (including the respective controls), a type of outcome assessed, and the method of outcome assessment.

Quality Assessment

The quality assessment of the articles included in this systematic review and meta-analysis was performed using the "Cambridge Quality Checklists" (13). In detail, it is composed of 3 checklists and evaluates the following domains: correlates, risk factors, and causal risk factors. The correlates checklist evaluates the appropriateness of the sample size and the quality of the outcome measurements. The checklist for risk factors assigns high-quality scores only to those studies with appropriate time-ordered data. Finally, the causal risk factor checklist evaluates the type of study design, assigning the highest score to randomized clinical trials and the lowest score to cross-sectional studies without a control group. To draw confident conclusions about correlates, the correlated score must be high. To draw confident conclusions about risk factors, both the checklists for correlates and risk factor scores must be high. To draw confident conclusions about causal risk factors, all scores on the 3-checklist scores must be high.

Statistical Analysis

Given the heterogeneity of outcome measurement methods, the standard mean difference (SMD) with the 95% confidential interval (CI) was calculated for quantitative variables (hormonal values, sperm parameters, testicular volume, and mean follicle count). Because age at menarche was assessed in all girls by questionnaire, the mean difference (MD) with the 95% CI was calculated for this outcome. Instead, the odds ratio (OR) with the 95% CI was calculated for dichotomous variables (Tanner stage, menarche prevalence, and polycystic ovary syndrome [PCOS] prevalence). The formula suggested by Hozo et al. (14), which allows estimating the mean and standard deviation having only the median and range values, was used for the studies whose data are reported as median and range, and it was not possible to have those values from the investigators. For hormone levels, all values were standardized to the same unit of measurement before being used in the quantitative analysis. In particular, for testosterone, all values were converted to nmol/L, for luteinizing hormone (LH) and follicle-stimulating hormone (FSH) to IU/L, for sex hormone-binding globulin (SHBG) to nmol/L, for inhibin B to pg/mL, for 17ß-estradiol to ng/L, for antimüllerian hormone (AMH) to ng/mL, and for dehydroepiandrosterone sulfate (DHEA) to μmol/L.

The Cochran-Q and I^2 statistics were used to evaluate the statistical heterogeneity. Specifically, if I^2 resulted in $\leq$50%, the studies' variation was considered homogenous. When I^2 was higher than 50%, there was significant heterogeneity between the studies. Because the population or a true effect can be expected to vary between studies, then a random effects model was used for the analysis. All P values of $\leq$.05 were considered statistically significant. Quantitative analysis

TABLE 1

Selection criteria in included studies Population, Exposure, Comparison/Comparator, Outcomes, Study design model of the current systematic review and meta-analysis.

PICOS criteria	Included	Excluded
Population	Human offspring	/
Exposure	ART (including IVF, ICSI, IUI, FET, ET, COS, OI, and HT)	Studies in which the offspring was born from infertile parents who did not receive fertility treatment
Comparison	SC	/
Outcome	Pubertal development for both sex Hormonal function for both sex Sperm function for male Follicular count for female	/
Study design	Observational cohort, cross-sectional, and case-control study designs	Case reports, comments, letters to the editor, systematic or narrative reviews, in vitro studies, and animal studies

ART = assisted reproductive techniques; COS = controlled ovarian stimulation; ET = embryo transfer; FET = frozen embryo transfer; HT = hormonal treatment; ICSI = intracytoplasmic sperm injection; IUI = intrauterine insemination; IVF = in vitro fertilization; OI = ovulation induction; SC = spontaneous conception.

Crafa. ART and offspring gonadal function. Fertil Steril Rev 2023.

was performed using RevMan software version 5.4 (Cochrane Collaboration, Oxford, United Kingdom) and Comprehensive Meta-Analysis Software (version 2) (Englewood, NJ).

We performed a cumulative analysis to evaluate the chronological trend of statistical significance over a period of time. To accomplish this, the effect size and the corresponding CI were calculated after chronologically adding each study. The *P* value trend and statistical inference were used to draw inferences about the strength of the association, its vulnerability to variation, and the history of variation.

We performed the sensitivity analysis with the exclusion method one study at a time. Therefore, the pooled effect size and corresponding CI were calculated after the exclusion of one study at a time. A study that resulted in an inference changing after its exclusion was labeled a "-sensitive study."

We analyzed the presence of publication bias quantitatively from the asymmetry of the funnel plot, which suggested some missing studies on one side of the graph. A quantitative analysis of publication bias was performed using Egger's intercept test, which assessed the statistical significance of publication bias. In cases of publication bias, unbiased estimates were calculated using the "trim and fill" method (15).

RESULTS

Using the search strategy mentioned above, we found 5,353 records. After excluding 2,496 duplicates, the remaining 2,857 articles were evaluated for inclusion in the systematic review. Of these, 1,234 were deemed not pertinent after having read their titles and abstracts; 1,605 were excluded because they were reviews (n = 580), systematic reviews and meta-analyses (n = 10), conference proceedings (n = 63), book chapters (n = 60), comments (n = 25), letters to the editor (n = 12), editorials (n = 11), short surveys (n = 13), and animal studies (n = 830). The remaining 19 articles have been carefully read. On the basis of the inclusion and exclusion criteria, 3 studies were excluded because of the inability to extract the data required, and one study was excluded because the offspring were born to infertile parents who did not receive infertility treatment to achieve pregnancy (16). Finally, 15 articles met our inclusion criteria. Therefore, they were included in this meta-analysis (17–31) (Fig. 1).

Information on the study design, the type of population and their age, the outcomes analyzed, and the methodology used to assess the outcome is reported in Table 2. Analysis of study quality showed that all studies had a medium risk of bias (Table 3).

Tanner Stage

Four studies evaluated the presence of differences in pubertal development as assessed by the Tanner stage between ART-conceived children and SC children (17–20). None of the studies found significant differences in the prevalence of different Tanner stages between the 2 groups regarding both pubic hair and genitalia in males and pubic hair and breasts in females. Only the study by Belva et al. (17) showed the presence of intracytoplasmic sperm injection (ICSI)-conceived girls with reduced breast development compared with the control group, even after adjustment for risk factors (age and body mass index, maternal educational level and birth weight, and gestational age and parity). Quantitative analysis was only possible for 2 of these 4 studies (17, 20). This analysis showed a comparable OR for specific Tanner genital stages between boys born using ICSI and these SC (Fig. 2). In particular, the OR of Tanner stage 2 did not significantly differ between the ICSI and SC groups (OR = 1.69 [0.60, 4.77]; *P*=.32) (Fig. 2A). Heterogeneity analysis showed the absence of heterogeneity, as demonstrated by the Q-test (Q-value = 0.021; *P*=.885) and $I^2 = 0$. Similarly, the OR of Tanner stage 3 did not significantly differ in ICSI vs. boys with SC (OR = 1.06 [0.43, 2.64]; *P*=.90) (Fig. 2B). In this case, a significant heterogeneity (Q-test: Q-value = 2.677; *P*=.102; $I^2 = 63\%$) was found among the studies. The OR of Tanner stage 4 did not differ in ICSI vs. boys with SC (OR = 0.88 [0.54, 1.44]; *P*=.62) (Fig. 2C). The analysis of heterogeneity (Q-test: Q-value = 1.688; *P*=.194; $I^2 = 41\%$) showed significant heterogeneity. The OR of Tanner stage 5 did not differ in ICSI vs. SC boys (OR = 0.89 [0.58, 1.35]; *P*=.57) (Fig. 2D). In this case, there was an absence of significant heterogeneity (Q-test: Q-value = 1.142; *P*=.707; $I^2 = 0$).

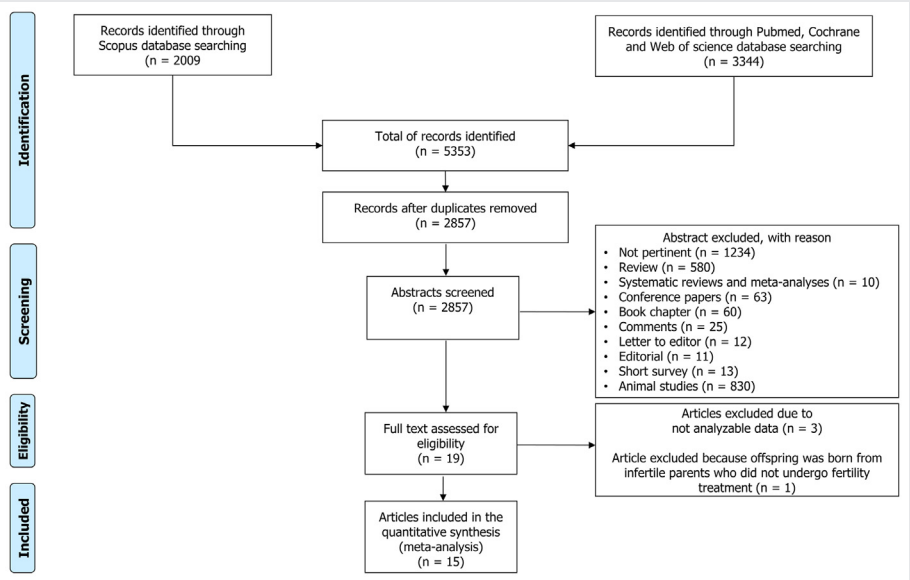
No study was sensitive enough to alter the above-reported results (data not shown). Cumulative analysis showed that these results remained nonsignificant after adding each study (data not shown). A publication bias analysis could not be performed because of the limited number of studies.

The OR of the Tanner stage 2 female breast in the ICSI vs. SC group was not statistically significant (OR = 2.32 [0.49, 10.99]; *P*=.29) (Fig. 3A), and no heterogeneity was found between studies (Q-test: Q-value = 0.572; *P*=.449; $I^2 = 0$). On the other hand, the quantitative analysis showed a slightly statistically higher OR of the Tanner stage 3 female breast in the ICSI group compared with the SC group (OR = 1.90 [1.04, 3.49]; *P*=.04) (Fig. 3B) with no interstudy heterogeneity (Q-test: Q-value = 0.1276; *P*=.59; $I^2 = 22\%$). One study (17) was sensitive enough to alter the above-reported results. Its removal resulted in the loss of significance of the OR (OR = 1.284 [0.528, 3.124]; *P*=.582). The ORs of Tanner stage 4 (OR= 0.75 [0.50, 1.12]; *P*=0.16) and 5 (OR = 0.70 [0.44, 1.12]; *P*=.16) (Fig. 3C, D, respectively) were nonsignificant. No heterogeneity was found between studies (Q-test: Q-value = 0.808; *P*=.352; $I^2 = 0$; Q-test: Q-value = 0.117; *P*=.732; $I^2 = 0$) for both analyses. No study was sensitive enough to alter the above-reported results (data not shown). The cumulative analysis showed that these results remained nonsignificant after the addition of each study (data not shown). An analysis of publication bias could not be performed because of the limited number of studies.

Testicular and Ovarian Volume

Data on testicular volume are present in 3 studies (21–23). In one study, testicular length was analyzed in 88 prepubertal children conceived through ICSI. The length of the testis and penis was normal for the age (21). The other 2 studies evaluated the testicular volume of young adults conceived with ART compared with that of young SC adults. Notably,

FIGURE 1



Flow chart of the studies included in the systematic review and meta-analysis.

Crafa. ART and offspring gonadal function. Fertil Steril Rev 2023.

Jensen et al. (22) found that young adults born to mothers who had undergone fertility treatments had, on average, lower testicular volume than SC young adults. This finding is not confirmed by the study by Belva et al. (23), which showed a mean testicular volume in boys born from ICSI comparable with that of SC boys. Quantitative analysis of these 2 studies showed no differences in testicular volume between groups conceived with ART and those after SC (SMD = −0.05 [−0.37, 0.27]; P=.75). Again, there was no heterogeneity between studies (Q-test: Q-value = 1.907; P=.167; I^2 = 47%) (Fig. 4).

Only 3 studies have evaluated ovarian volume using ultrasound. In one study, the analysis was performed on newborns with breast buttons born from ART. The study found prepubertal ovarian volume consistent with age (24). In a second study, the ovaries were analyzed in pubertal-age girls conceived using ICSI and compared with pubertal SC girls. Volumes did not differ significantly between the 2 groups. However, the data are not shown (25). Furthermore, the third study, which compared ART-conceived pubertal girls with SC peers, failed to find a difference in ovarian volume between the 2 groups (26).

Two studies analyzed follicular counts in girls born from ART vs. SC (25, 26). Quantitative analysis observed no differences in mean follicular count between the 2 groups (SMD = −0.07 [−0.47, 0.34]; P=.75), and no interstudy heterogeneity was found (Q-test: Q-value = 0.639; P=.424; I^2 = 0) (Fig. 5A). Furthermore, in both groups, the proportion of women with polycystic ovarian appearance was comparable (OR = 1.03 [0.12, 8.91]; P=.98). No interstudy heterogeneity was found (Q-test: Q-value = 1.711; P=.191; I^2 = 44%) (Fig. 5B).

For both testicular and ovarian volume analyses, no studies were sensitive enough to alter the above results (data not shown). Cumulative analysis showed that these results remained nonsignificant after adding each study (data not shown). A publication bias analysis could not be performed because of the limited number of studies.

Pubertal Milestones

The achievement of important pubertal milestones was evaluated in 6 studies. These include age at menarche in girls, age of facial hair appearance, voice deepening and first nocturnal ejaculation in boys, and the appearance of axillary hair and acne in both genders (17–20, 26, 27). All studies except one that evaluated age at menarche showed no difference between children conceived using ART and SC children (17, 18, 20). The only study that found a difference reported a delay in age at menarche in the ART group compared with the control group, even after correction for confounding factors (26). The quantitative analysis of the 4 studies that evaluated this aspect did not confirm any difference in the mean age at menarche between girls conceived using ART or SC girls regardless of the ART method used (MD = 0.06 [−0.24, 0.35]; P=.70). Heterogeneity analysis showed no heterogeneity between studies (Q-test: Q-value = 8.378; P=.039; I^2 = 49%) (Fig. 6A). Also, Egger's regression model and funnel plots reported no risk of bias (intercept = 3.16, 95% CI: −9.98–16.29, P=.20) (Supplemental Fig. 1, available online). Similarly, the OR of having menarche was also comparable between the group of girls conceived using ART and SC girls, regardless of the technique used (OR = 0.97 [0.67, 1.41]; P=.88). No interstudy

TABLE 2

The Main characteristics of the studies included.

Author	Study design	Country	Data collection period	Case sample size	Control sample size	Age of cases	Age of controls	Outcomes assessed	Method of evaluation of the outcome
Belva et al., 2010 (31)	Longitudinal cohort study	Belgium	Between January 2008 and November 2009	50 boys were conceived using ICSI	/	14 y	/	Inhibin B	ELISA
Belva et al., 2011 (30)	Cross-sectional	Belgium	Between January 2008 and April 2010	58 boys were conceived using ICSI	62 boys SC	14.3 ± 0.4 y	14.1 ± 0.4 y	Salivary testosterone	Radioimmunoassay
Belva et al., 2012 (17)	Cross-sectional	Belgium	Between January 2008 and March 2011	217 ICSI-conceived children (F = 101; M = 116)	223 SC children (F = 108; M = 115)	F = 14 ± 0.5 y M = 14 ± 0.4 y	F = 14.3 ± 0.3 y M = 14.3 ± 0.3 y	Tanner scores for breast and pubic hair in females and for genital and pubic hair in males	Physical examination
								The mean age of menarche	Questionnaire
Belva et al., 2016 (23)	Cross-sectional	Belgium	Between March 2013 and April 2016	54 young male adults conceived using ICSI	57 SC young male adults	19.5 ± 0.7 y	20 ± 1.2 y	Sperm parameters	WHO 2010 manual
Belva et al., 2017a (25)	Cross-sectional	Belgium	Between March 2013 to April 2016	71 singleton females conceived using ICSI	81 SC women age-matched	19.3 ± 0.6 y	19.5 ± 0.9 y	AMH, Estradiol, FSH, and DHEAs	Electro-chemiluminescent assay
								Mean follicle number per ovary	Transvaginal ultrasound
Belva et al., 2017b (29)	Cross-sectional	Belgium	Between March 2013 to April 2016	54 young male adults conceived using ICSI	57 SC young male adults	19.5 ± 0.7 y	20 ± 1.2 y	FSH, LH, and TT	Electro-chemiluminescence immunoassay
								Inhibin B	ELISA
Beydoun et al., 2011 (27)	Cross-sectional	USA	NR	166 (F = 95; M = 71) young adults conceived using IVF	/	TS = 21.2 ± 2.2 y	/	Puberty onset and age of achievement of pubertal developmental milestones	Self-administered questionnaire
Ceelen et al., 2008 (18)	Cross-sectional	Netherlands	Between March 2003 and March 2006	233 IVF children (F = 118; M = 115)	233 SC children from infertile mothers (F = 118; M = 115)	F = 12.4 ± 2.6 y M = 12 ± 2.6 y	F = 12.4 ± 2.6 y M = 12 ± 2.6 y	Pubertal development	Tanner's stage was assessed using a physical examination
				49 IVF girls	49 SC girls from infertile mothers	14.9 ± 1.5 y	14.8 ± 1.7 y	The mean age of menarche	Questionnaire
				67 IVF boys and 72 IVF girls	68 boys and 75 girls SC from infertile mothers	F = 13.7 ± 2.1 y M = 13.6 ± 2 y	F = 13.5 ± 2.1 y M = 13.6 ± 2 y	Skeletal age	RX
				40 IVF boys and 19 IVF girls	35 boys and 20 girls SC from infertile mothers	F = 11.4 ± 0.9 y M = 14.7 ± 1.6 y	F = 11.2 ± 0.9 y M = 15 ± 1.5 y	LH, FSH	Immunofluorometric assays
								Estradiol	Double-antibody Radioimmunoassay
								TT and DHEAS	Coat-A-Count radioimmunoassay
De Schepper et al., 2009 (21)	Cross-sectional	Belgium	NR	88 male children conceived using ICSI	/	8.5 (7.9–9.1) y median (min–max)	/	Testicular size expressed the length in mm and penis size	Physical examination
				59 male children born using ICSI	37 SC male children	8.5 (7.9–9.1) y median (min–max)	Aged between 7.8–9.7 y	Inhibin B	ELISA
								AMH	Two-step two-site immunoassay

Crafa. ART and offspring gonadal function. Fertil Steril Rev 2023.

TABLE 2

Continued.

Author	Study design	Country	Data collection period	Case sample size	Control sample size	Age of cases	Age of controls	Outcomes assessed	Method of evaluation of the outcome
Ernst et al., 2019 (19)	Prospective cohort study	Denmark	From 2012 to 2018	1464 children conceived using ART (902 using COS, 480 using IVF or ICSI, 82 others) (F = 745; M = 719)	8,541 SC children from fertile parents (F = 4,403; M = 4,138) 3,280 SC children from infertile parents (F = 1,669; M = 1,611)	/	/	Puberty onset and age of achievement of pubertal developmental milestones	Self-administered questionnaire
Jensen et al., 2007 (22)	Cross-sectional	Denmark	From 2001 to 2005	47 males born from mothers who received fertility treatment (25/47 received hormonal treatment)	1,702 males born from a fertile mother	/	/	Sperm parameters	WHO 1992 was modified in accordance with Jorgensen et al. 1997 (48)
								Rate of cryptorchidism	Question
								Testicular volume	Physical examination
				39 males born from mothers who received fertility treatment	1,222 males born from a fertile mother	/	/	FSH, LH, and TT	Time-resolved-immunofluorometric assay
								Inhibin B	Enzymatic assay
								SHBG	/
Kai et al., 2007 (28)	Cross-sectional	Denmark	Between 1998 and 2005	72 male ICSI newborns and 83 male IVF newborns	598 SC male newborns	ICSI = 3.5 (2.6–4.5) mo IVF = 3.4 (2.6–4.5) mo	3 (2.5–4.2) mo	Testosterone and estradiol	RIA
								fT	Vermeulen's formula
								LH, FSH, and SHBG	Fluoroimmunometric assay
								Inhibin B	Enzyme-immunometric assay
Merino et al., 2019 (26)	Cross-sectional	Chile	NR	22 girls conceived using ART	53 SC girls	14.3 ± 1.5 y	13.9 ± 1.3 y	Follicle count and ovarian size	Ultrasound
								Inhibin B and AMH	Enzymatic Assay
								Testosterone and androstenedione	Radioimmunoassay
								LH, FSH, and SHBG	Immunoradiometric assay
Rojas-Marcos et al., 2005 (24)	Case series	USA	Between May 2001 and April 2004	7 children conceived using ART (F = 6; M = 1)	/	11.6 ± 5.9 mo	/	Cortisol, Estradiol, DHEAs, and TT	Chemiluminescence
								17αOH-P, D4A, and estrone	RIA
								LH and FSH	Microenzyme immunoassay
								TT, 17αOH-P, and DHEAs	RIA for patients 2,4,6, and 7
								Ovarian size	Ultrasound
								Bone age	Radiography
Sonntag et al., 2020 (20)	Cross-sectional	Germany	NR	274 ICSI singleton (F = 141; M = 133)	273 SC (F = 142; M = 131)	TS = 16.5 ± 0.5 y F = 16.5 ± 0.5 y M = 16.5 ± 0.5 y	TS = 16.5 ± 0.8 y F = 16.5 ± 0.8 y M = 16.3 ± 0.9 y	Pubertal development	Tanner's stage was assessed using a physical examination
								FSH, LH, estradiol, TT, and AMH	Electro-chemiluminescence immunoassay
								Inhibin B	ELISA

AMH = antimüllerian hormone; DHEAs = dehydroepiandrosterone sulfate; ELISA = enzyme-linked immunosorbent assay; F = female; FSH = follicle-stimulating hormone; fT = free testosterone; ICSI = intracytoplasmic sperm injection; IVF = in vitro fertilization; LH = luteinizing hormone; M = male; NR = not reported SC = spontaneously conceived; RIA = radioimmunoassay; TT = total testosterone; 17αOH-P = 17α-hydroxyprogesterone; SHBG = sex hormone-binding globulin; / = data not available.

Crafa. ART and offspring gonadal function. Fertil Steril Rev 2023.

TABLE 3

Evaluation of studies quality using "The Cambridge Quality Checklists."

Author and year of publication	Checklist for correlates	Checklists of risk factors	Checklist of causal risk factors	Total
Belva et al., 2010 (31)	2	3	3	8/15
Belva et al., 2011 (30)	2	1	5	8/15
Belva et al., 2012 (17)	2	1	5	8/15
Belva et al., 2016 (23)	3	1	5	9/15
Belva et al., 2017a (25)	1	1	5	7/15
Belva et al., 2017b (29)	2	1	5	8/15
Beydoun et al., 2011 (27)	1	1	1	3/15
Ceelen et al., 2008 (18)	1	1	2	4/15
De Schepper et al., 2009 (21)	1	1	2	4/15
Ernst et al., 2019 (19)	2	1	5	8/15
Jensen et al., 2006 (22)	3	1	2	6/15
Kai et al., 2007 (28)	2	1	5	8/15
Merino et al., 2019 (26)	2	1	5	8/15
Rojas-Marcos et al., 2005 (24)	1	3	3	7/15
Sonntag et al., 2020 (20)	2	1	5	8/15

Crafa. ART and offspring gonadal function. Fertil Steril Rev 2023.

heterogeneity was found (Q-test: Q-value = 0.045; P=.832; I^2 = 0) (Fig. 6B). Furthermore, only one study observed a lower ovulation rate in girls conceived using ART compared with SC girls (26).

The few studies that have evaluated the achievement of pubertal milestones in boys have found no differences between the ART-conceived and the SC groups (19, 27).

For the analysis of pubertal milestones, no studies were sensitive enough to alter the above-reported results (data not shown). A cumulative analysis showed that these results remained nonsignificant after adding each study (data not shown).

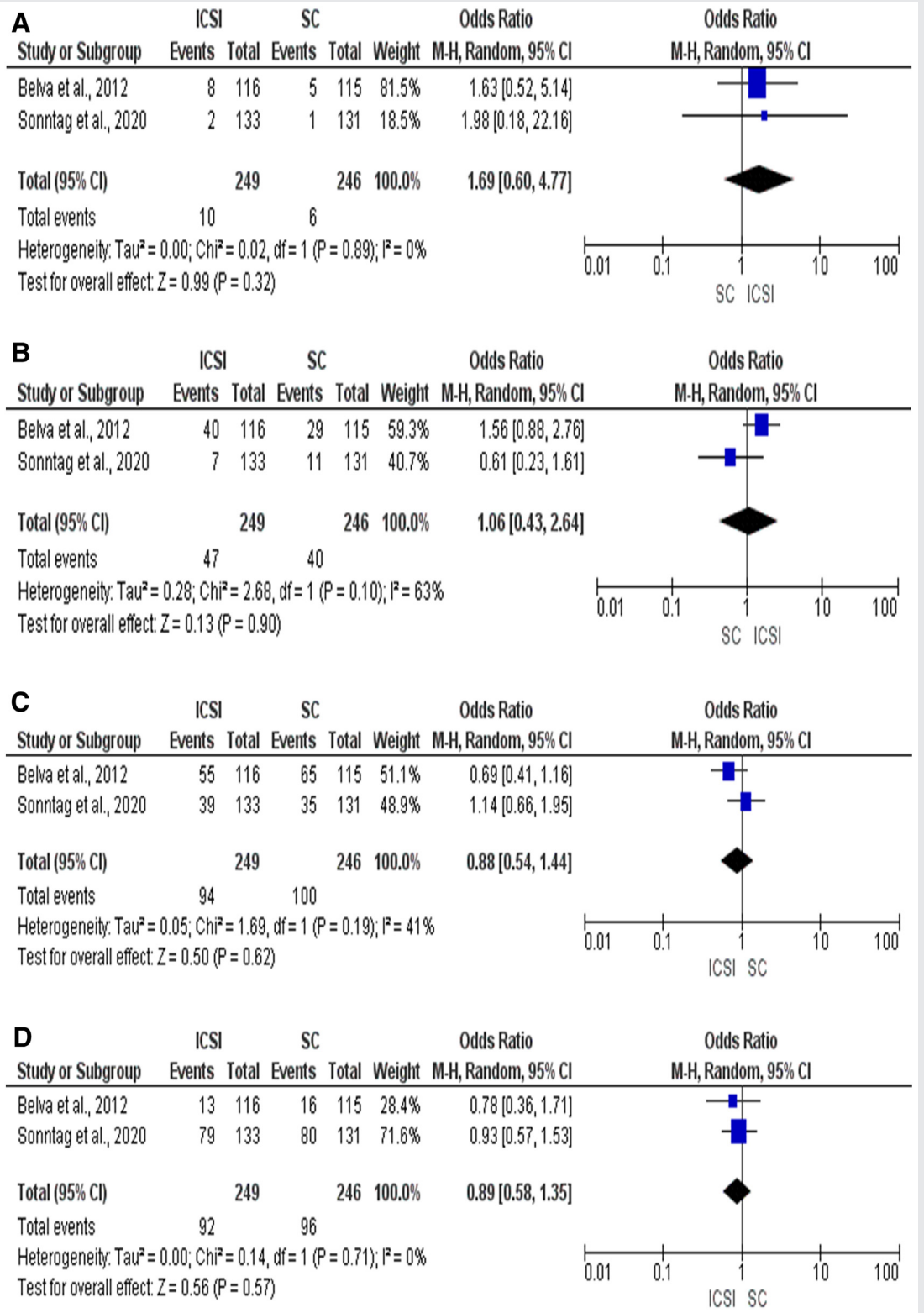
Hormonal Values

Eleven of the 15 studies identified in this meta-analysis evaluated serum hormone levels in ART-conceived children (18, 20–22, 24–26, 28–31), and in 9 of these studies, there is a comparison with a population of SC children (18, 20–22, 25, 26, 28–30). In 2 studies, hormone measurements were performed on newborns (23, 27). In particular, Mau Kai et al. (28) evaluated hormone levels in children with an average age of 3 months, the period when mini-puberty is present, and showed that children born from ART have lower total and free testosterone values and a higher LH/testosterone ratio than SC children even after adequate for gestational age and singleton selection. This finding was confirmed only in ICSI-conceived children performed for paternal infertility. The children had a 27.3% and 23.4% reduction in total and free testosterone levels, respectively, and a 60% increase in the LH/total testosterone ratio compared with SC children. In contrast, no differences were found between infants born from IVF performed for maternal infertility and SC infants. This finding seems to point attention to the possible role of paternal infertility in causing reproductive harm to offspring (28).

Another study compared the serum levels of AMH and inhibin B in 8-year-old prepubertal children born from ICSI with those of SC children of the same age. The study found no differences between the 2 groups. Furthermore, no differences were found in the levels of these 2 hormones in children conceived with ICSI children from fathers with severe oligozoospermia (concentration <5 mil/mL) compared with those with less severe forms (21).

Regarding hormone levels in pubertal ART-conceived boys, only one study analyzed salivary testosterone levels in ICSI-born and SC boys and found no differences between the 2 groups (30). Furthermore, no correlation was found between the severity of the father's oligozoospermia and the salivary testosterone levels in the offspring, thus also ruling out the role of paternal infertility (30). On the other hand, 4 studies evaluated the hormone levels in the blood (18, 20, 22, 29). Three of these studies found no statistically significant differences in hormone levels analyzed between the ART-conceived and SC boys (18, 22, 29). Although in the study by Jensen et al. (22), there was a trend toward a reduction in testosterone levels and the free androgen index (FAI). In contrast, in the study by Belva et. al. (29), there was a higher prevalence of boys with inhibin B<10 percentile and FSH of >90 percentile in the ICSI-conceived boys compared with SC boys. Conversely, the study by Sonntag et al. (20) showed a significant increase in estrogen and a decrease in the testosterone/17ß-estradiol ratio in boys conceived using ICSI compared with the SC group, and a trend toward lower levels of inhibin B. Quantitative analysis showed significantly reduced levels of SHBG (SMD = −0.24 [−0.47, −0.01]; P=.05] in ART patients than in controls, in particular in ICSI patients (Fig. 7A) and lower levels of inhibin B only in the ICSI subgroups (SMD = −0.24 [−0.44, −0.04]; P=.02] (Fig. 7B) (20, 22). In both analyses, no interstudy heterogeneity was found (Q-test: Q-value = 1.419; P=.234; I^2 = 29%, and Q-test: Q-value = 3.208; P=.201; I^2 = 37%, respectively). Furthermore, it revealed a trend toward a reduction in

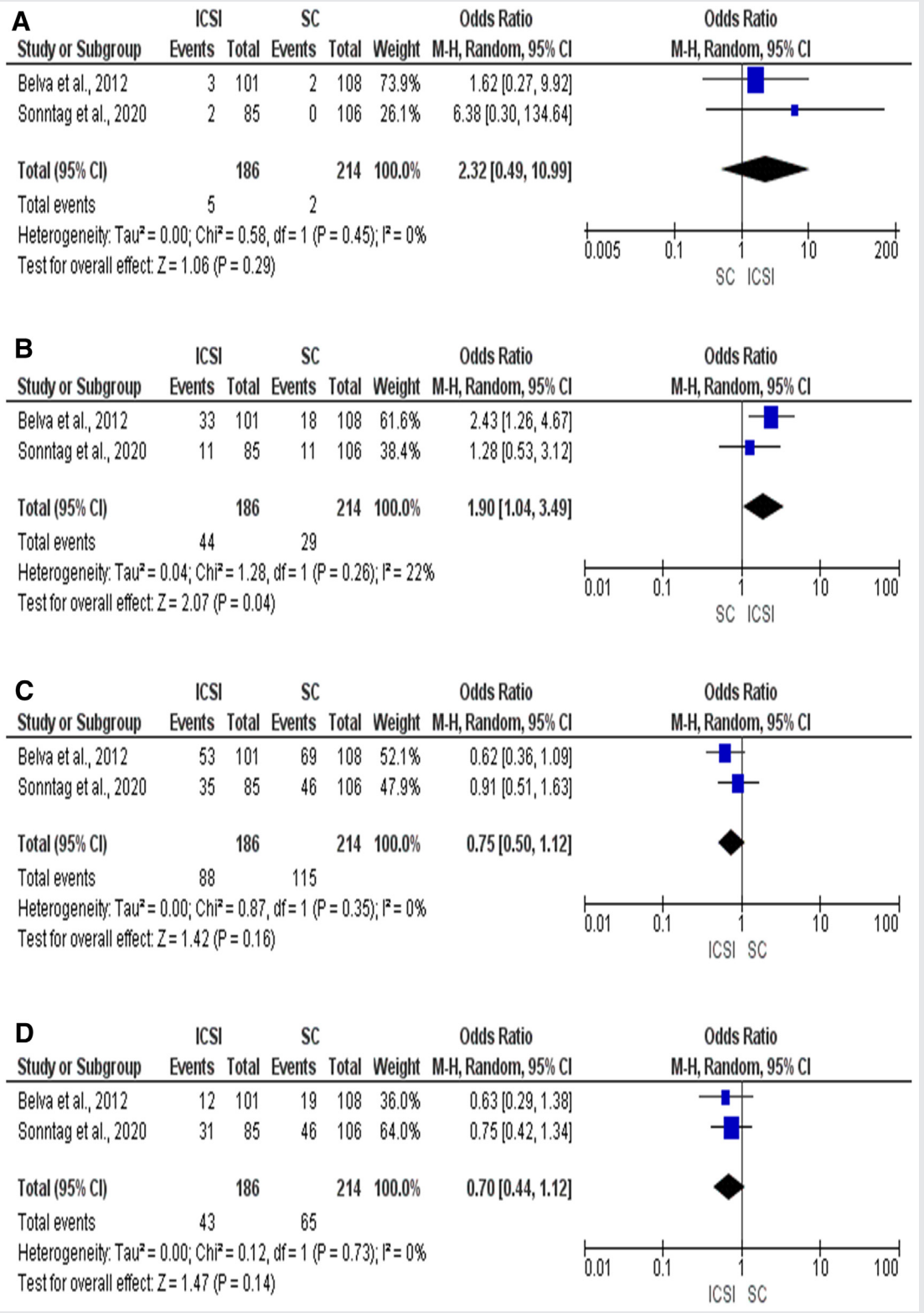
FIGURE 2



(**A**) Forest plot showing the odds ratio (OR) of Tanner stage 2 in male assisted reproductive technique (ART)-conceived children and spontaneously conceived (SC) children. (**B**) Forest plot showing the OR of Tanner stage 3 in male ART-conceived children and SC children. (**C**) Forest plot showing the OR of Tanner stage 4 in male ART-conceived and SC children. (**D**) Forest plot showing the OR of Tanner stage 5 in male ART-conceived children and SC children.

Crafa. ART and offspring gonadal function. Fertil Steril Rev 2023.

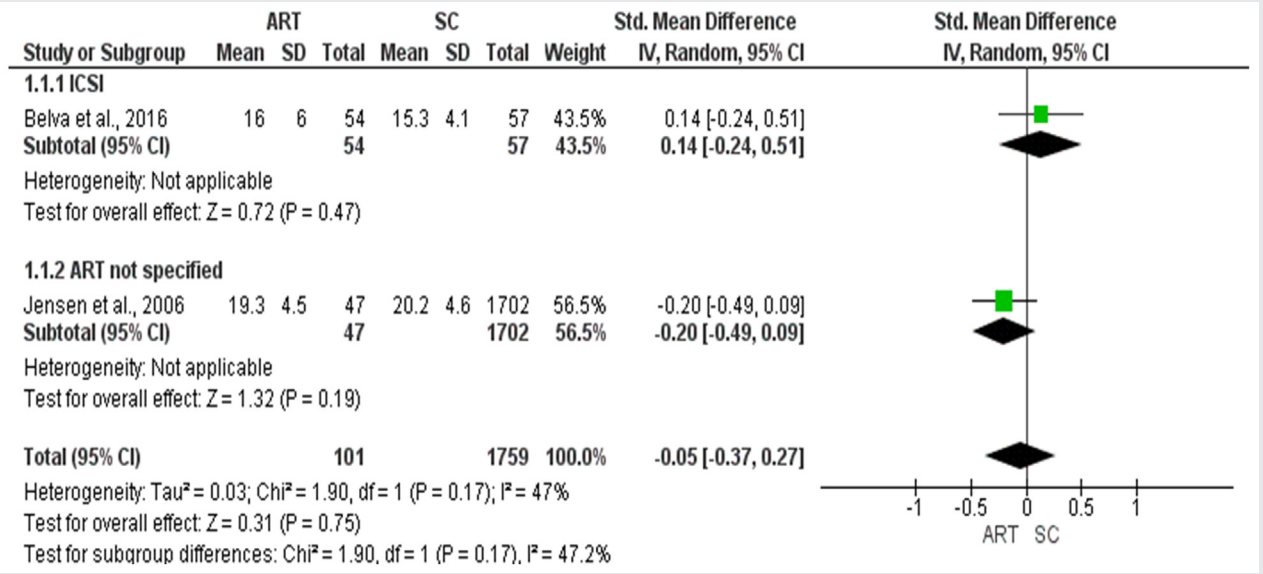
FIGURE 3



(**A**) Forest plot showing the odds ratio (OR) of Tanner stage 2 in female assisted reproductive technique (ART)-conceived children and spontaneously conceived (SC) children. (**B**) Forest plot showing the OR of Tanner stage 3 in female ART-conceived children and SC children. (**C**) Forest plot showing the OR of Tanner stage 4 in female ART-conceived children and SC children. (**D**) Forest plot showing the OR of Tanner stage 2 in female ART-conceived children and SC children.

Crafa. ART and offspring gonadal function. Fertil Steril Rev 2023.

FIGURE 4



Forest plot showing standardized mean difference (SMD) in testicular volume between assisted reproductive technique (ART)-conceived and spontaneously conceived (SC) boys.

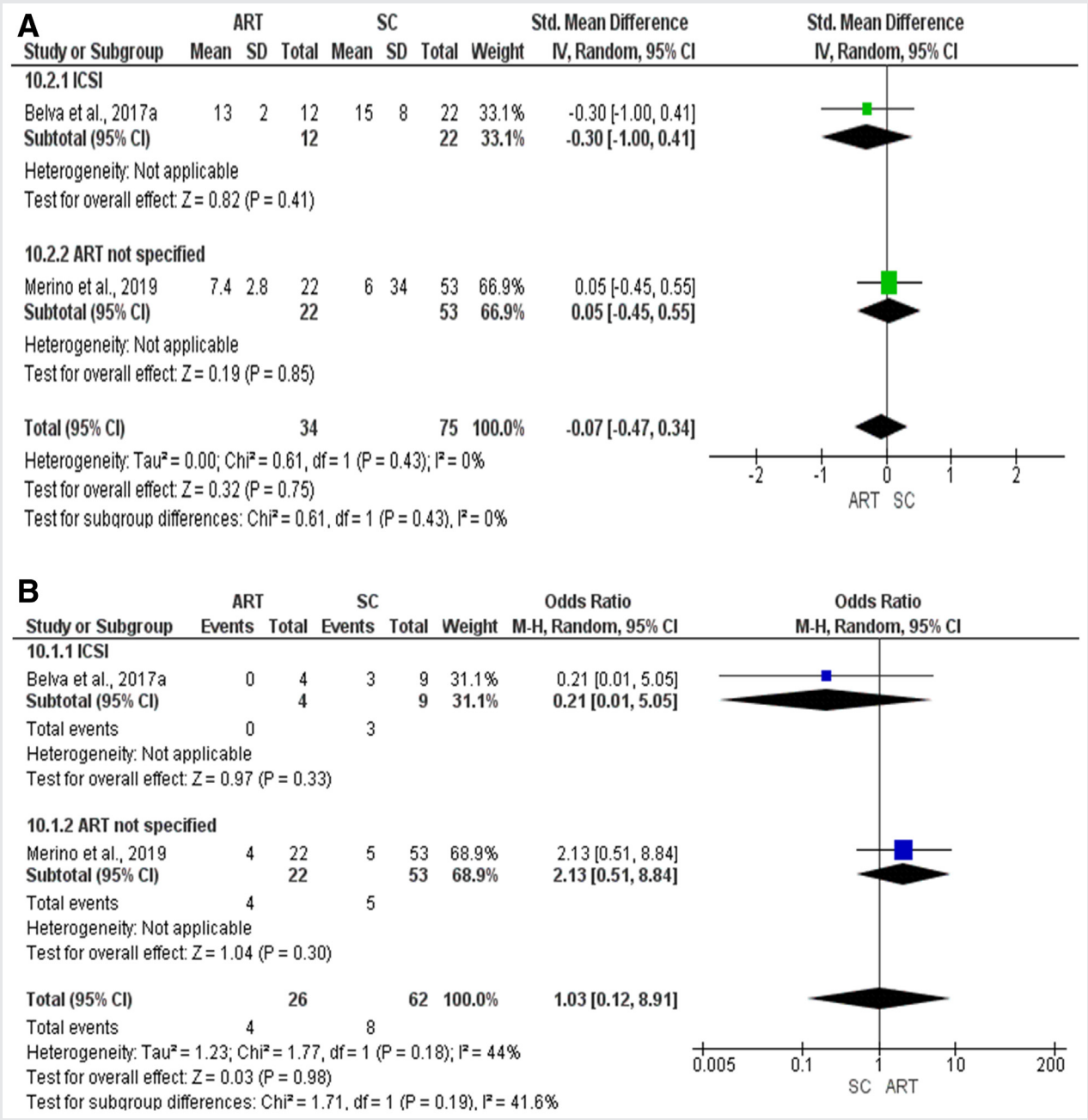
Crafa. ART and offspring gonadal function. Fertil Steril Rev 2023.

testosterone levels in boys conceived with ART compared with SC boys regardless of the ART method used (SMD= −0.16 [−0.32, 0.01]; *P*=.06] (Fig. 7C) (18, 20, 22, 29). No heterogeneity was found between studies (Q-test: Q-value = 0.289; *P*=.865; $I^2 = 0$).

Instead, not significantly different levels of LH (SMD = −0.11 [−0.27, 0.05]; *P*=.17] (Fig. 8A), FSH (SMD = 0.06 [−0.10, 0.22]; *P*=.47; $I^2 = 0$] (Fig. 8B), and FAI (SMD = 0.07 [−0.13, 0.27]; *P*=.49] (Fig. 8C) were found (18, 20, 22, 29). No heterogeneity was found between studies (Q-test: Q-value = 2.091; *P*=.554; $I^2 = 0$; and Q-test: Q-value =0.667; *P*=.881; $I^2 = 0$; and Q-test: Q-value = 1.097; *P*=.295; $I^2 = 8\%$). The Egger's regression model and funnel plots reported no risk of bias (intercept = −2.66, 95% CI: −7.72 to −2.41, *P*=.07, Supplemental Fig. 2A, available online: intercept = −0.38, 95% CI: −5.68 to 4.92, *P*=.39, Supplemental Fig. 2B). A publication bias analysis could not be performed for FAI because of the limited number of studies.

Regarding the female gender, only 5 studies reported hormonal values (18, 20, 24–26). One of these was performed on newborn girls with breast buttons and revealed normal prepubertal hormone levels, excluding the possibility of precocious puberty (24). Another 4 studies measured hormone levels in pubertal ART-conceived girls and compared them with SC girls (18, 20, 25, 26). Quantitative analysis was possible for these 4 studies. In particular, Sonntag et al. (20) included only women who were not using contraceptives and had 17ß-estradiol levels <60 pg/mL. Belva et al. (25) included only data from women who did not use contraceptives (25). The analysis found higher mean LH levels in ART-conceived girls than SC girls (SMD = 0.34 [0.04, 0.64]; *P*=.02) (Fig. 9A) (18, 20, 25, 26) and without heterogeneity between studies (Q-test: Q-value = 3.763; *P*=.288; $I^2 = 17\%$). The Egger's regression model and funnel plots reported no risk of bias (intercept = 2.20; 95% CI: −9.73– 4.14; *P*=.25, Supplemental Fig. 3, available online). The studies by Ceelen et al. (18) and Merino et al. (26) were identified as sensitive enough to change the results of the analysis (SMD = 0.25 [-0.04, 0.15]; *P*=.08 and SMD = 0.25 [−0.06, 0.16]; *P*=.11, respectively). No differences were found in DHEAs levels (SMD = 0.50 [−0.29, 1.30]; *P*=.22) in the group of girls conceived using ART compared with SC girls (18, 25). In contrast, Ceelen et al. (18) found higher levels of this hormone in girls born using IVF compared with SC controls (Fig. 9B) (17). Analysis of heterogeneity showed the presence of inter-study heterogeneity (Q-test: Q-value = 2.836; *P*=.09; $I^2 = 63\%$). No study was found to be sensitive enough to change these results. A cumulative analysis showed that data became nonsignificant after adding the study by Ceelen et al. (18) (data not shown). A publication bias analysis could not be performed because of the limited number of studies. Furthermore, no differences were found for FSH (SMD = 0.19 [−0.07, 0.45]; *P*=.16] and no heterogeneity was found between studies (Q-test: Q-value = 0.588; *P*=.899; $I^2 = 0$) (Fig. 10A). No study was sensitive enough to alter the above-reported results (data not shown). Furthermore, the cumulative analysis showed that these results remained nonsignificant after adding each study (data not shown). The Egger's regression model and funnel plots reported no risk of bias (intercept = 0.295; 95% CI: −4.63–5.23, *P*=.41, Supplemental Fig. 4A, available online).

FIGURE 5



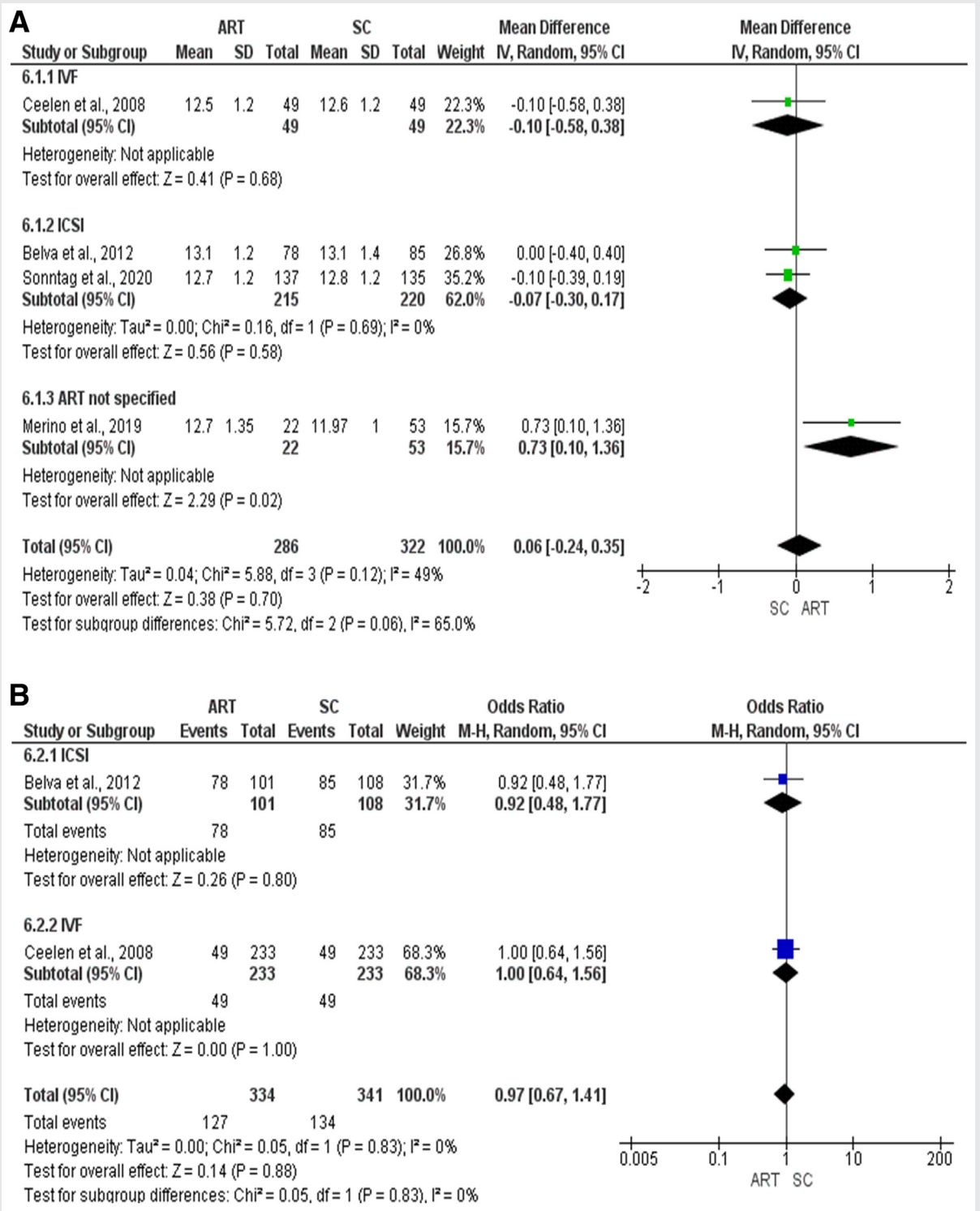
(**A**) Forest plot showing standardized mean difference (SMD) of follicular count between assisted reproductive technique (ART)-conceived and spontaneously conceived (SC) girls. (**B**) Forest plot showing the odds ratio (OR) of having polycystic ovaries in ART-conceived and SC girls.

Crafa. ART and offspring gonadal function. Fertil Steril Rev 2023.

Analysis of AMH showed the absence of interstudy heterogeneity (Q-test: Q-value = 2.535; P=.282; I^2 =19%), and resulted in the absence of any significant difference (SMD = −0.08 ([0.41, 0.25]; P=.62) (Fig. 10B). The Egger's regression model and funnel plots reported no risk of bias (intercept = −2.24; 95% CI: −49.52–45.04; P=.33, Supplemental Fig. 4B).

The serum levels of testosterone were not significantly different [SMD = −0.19 [−0.61, 0.23]; P=.38; (Q-test: Q-value = 1.753; P=.185; I^2 = 42%) (Fig. 10C), as well as those of the SHBG (SMD = 0.25 [−0.07, 0.56]; P=.12; Q-test: Q-value = 0.154; P=.694; I^2 = 0) (Fig. 10D), FAI (SMD = −0.27 [−0.59, 0.04]; P=.09; Q-test: Q-value =

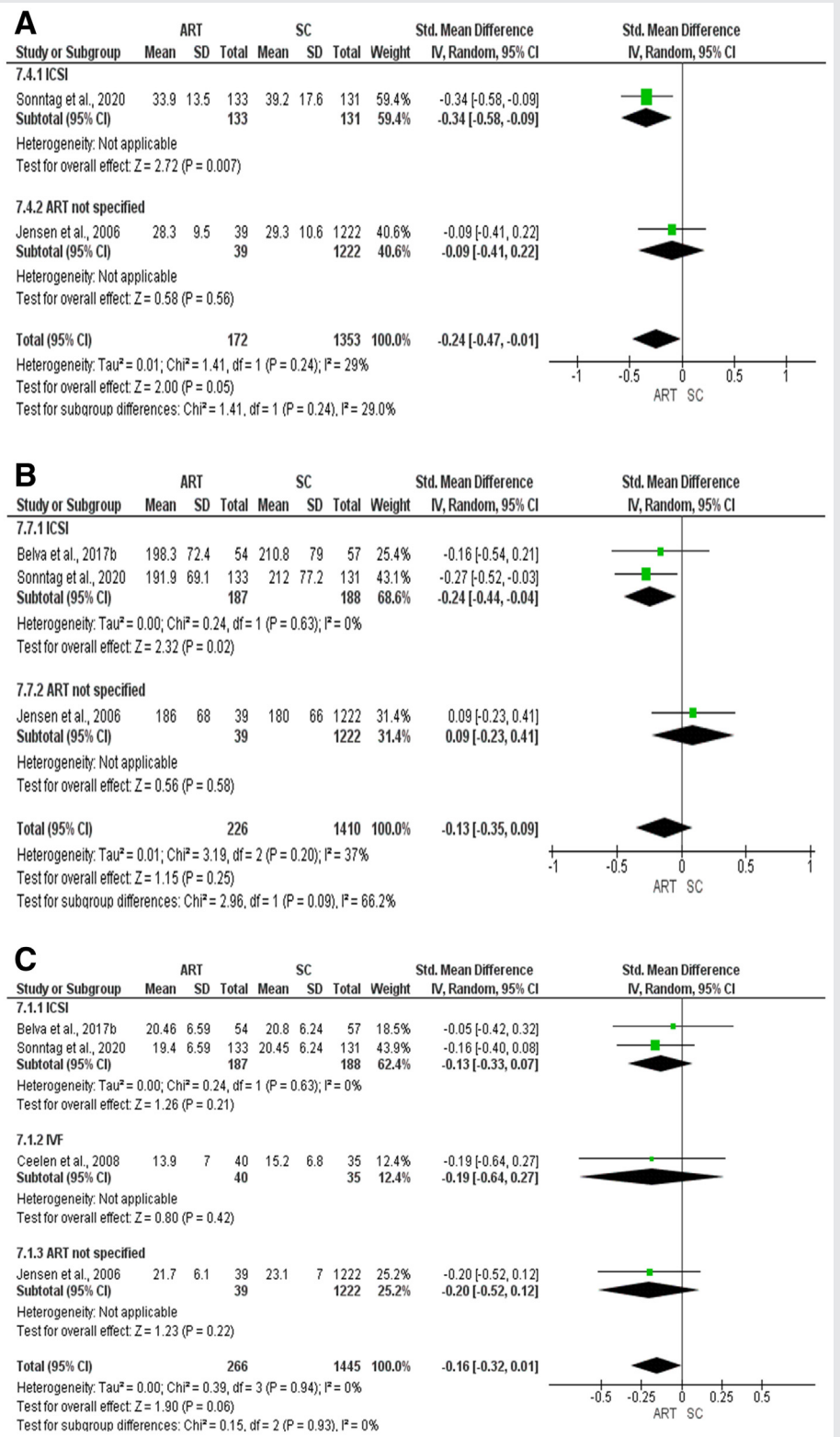
FIGURE 6



(**A**) Forest plot showing mean difference (MD) in age at menarche between assisted reproductive technique (ART)-conceived and spontaneously conceived (SC) girls. (**B**) Forest plot showing the odds ratio (OR) of having menarche in ART-conceived and SC girls.

Crafa. ART and offspring gonadal function. Fertil Steril Rev 2023.

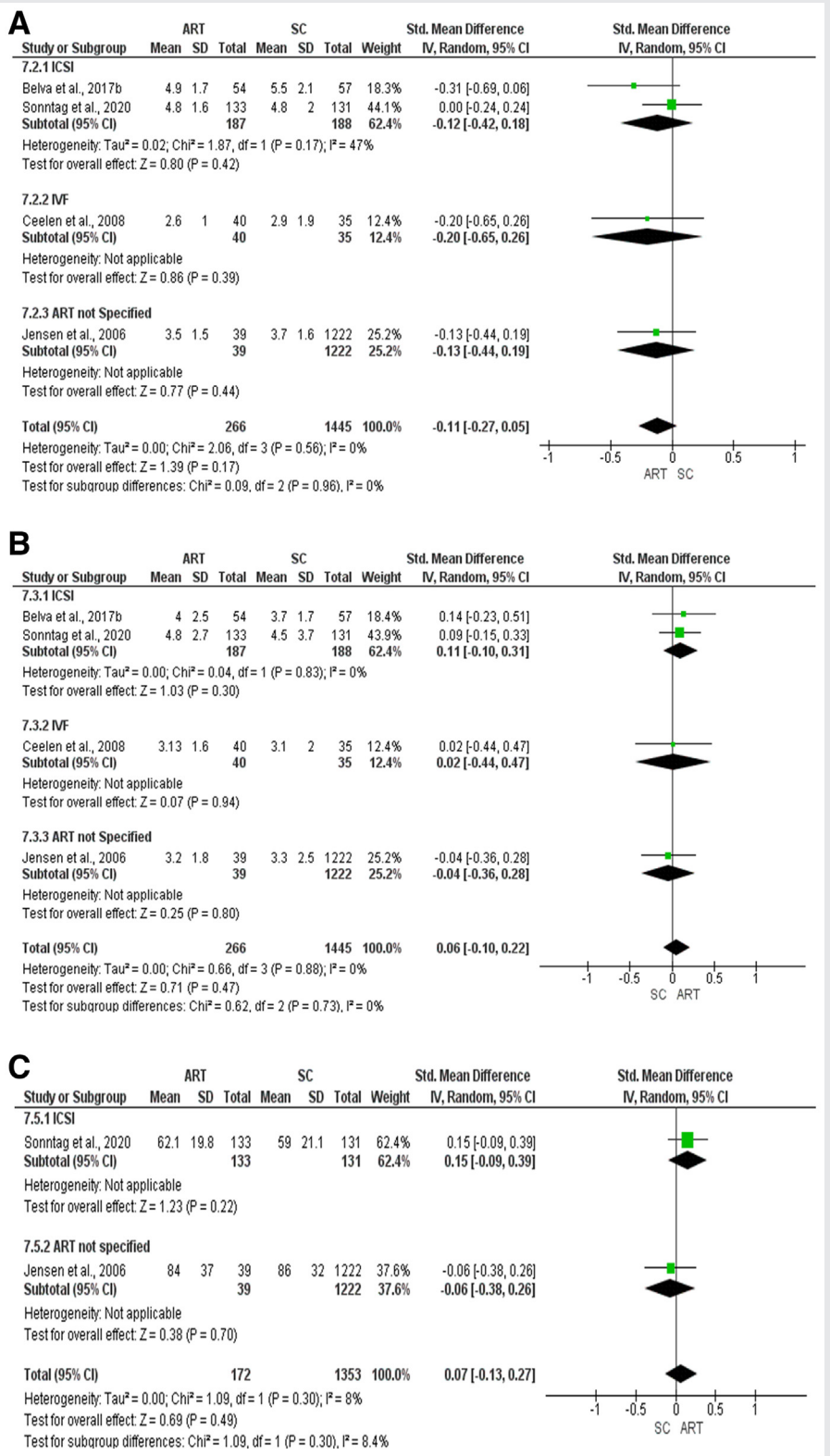
FIGURE 7



(**A**) Forest plot showing standardized mean difference (SMD) in sex hormone-binding globulin (SHBG) levels between assisted reproductive technique (ART)-conceived and spontaneously conceived (SC) boys. (**B**) Forest plot showing SMD in inhibin B levels between ART-conceived and SC boys. (**C**) Forest plot showing SMD in total testosterone (TT) levels between ART-conceived and SC boys.

Crafa. ART and offspring gonadal function. Fertil Steril Rev 2023.

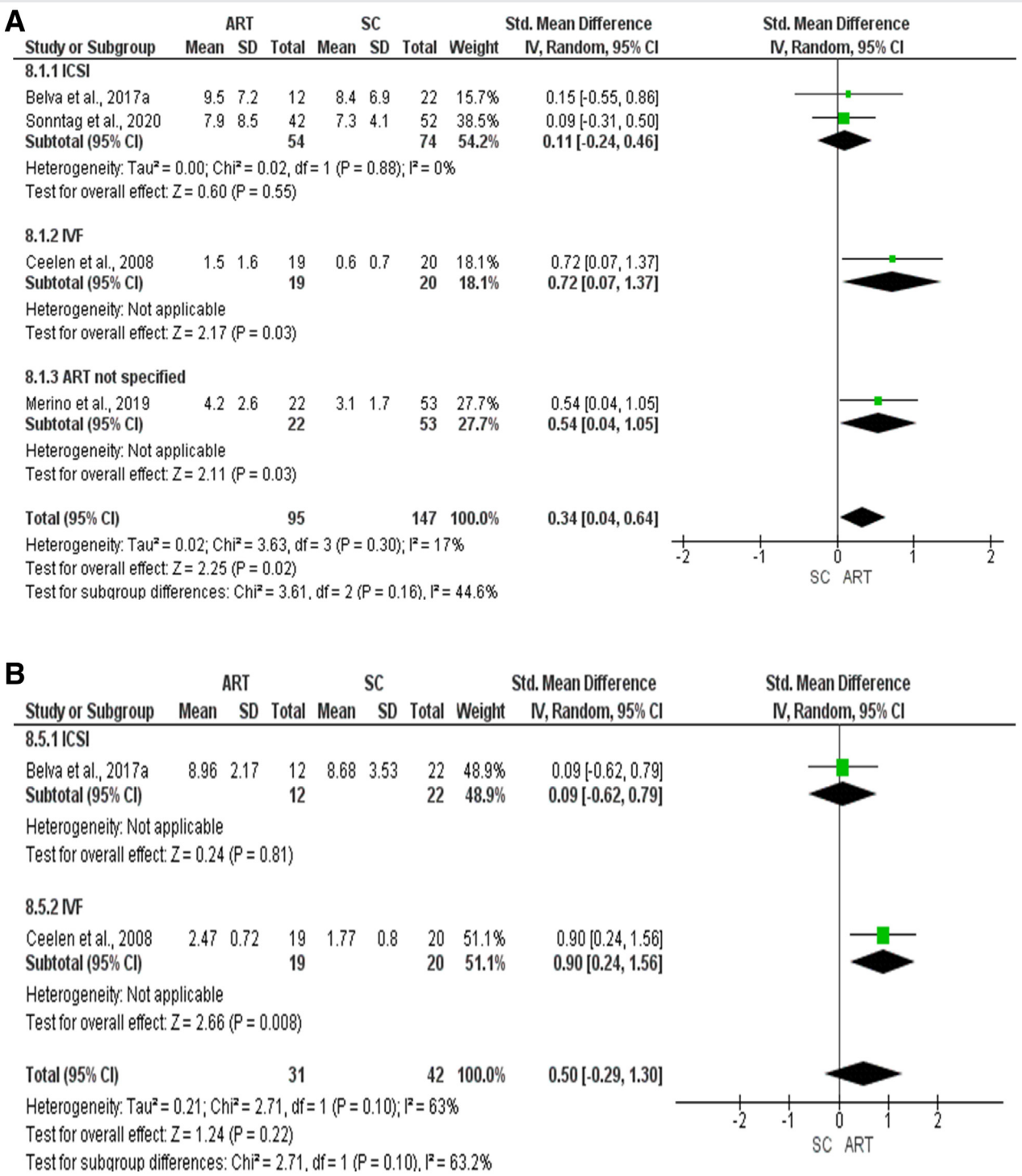
FIGURE 8



(**A**) Forest plot showing standardized mean difference (SMD) in luteinizing hormone (LH) levels between assisted reproductive technique (ART)-conceived and spontaneously conceived (SC) boys. (**B**) Forest plot showing SMD in follicle-stimulating hormone (FSH) levels between ART-conceived and SC boys. (**C**) Forest plot showing SMD in free androgen index (FAI) between ART-conceived and SC boys.

Crafa. ART and offspring gonadal function. Fertil Steril Rev 2023.

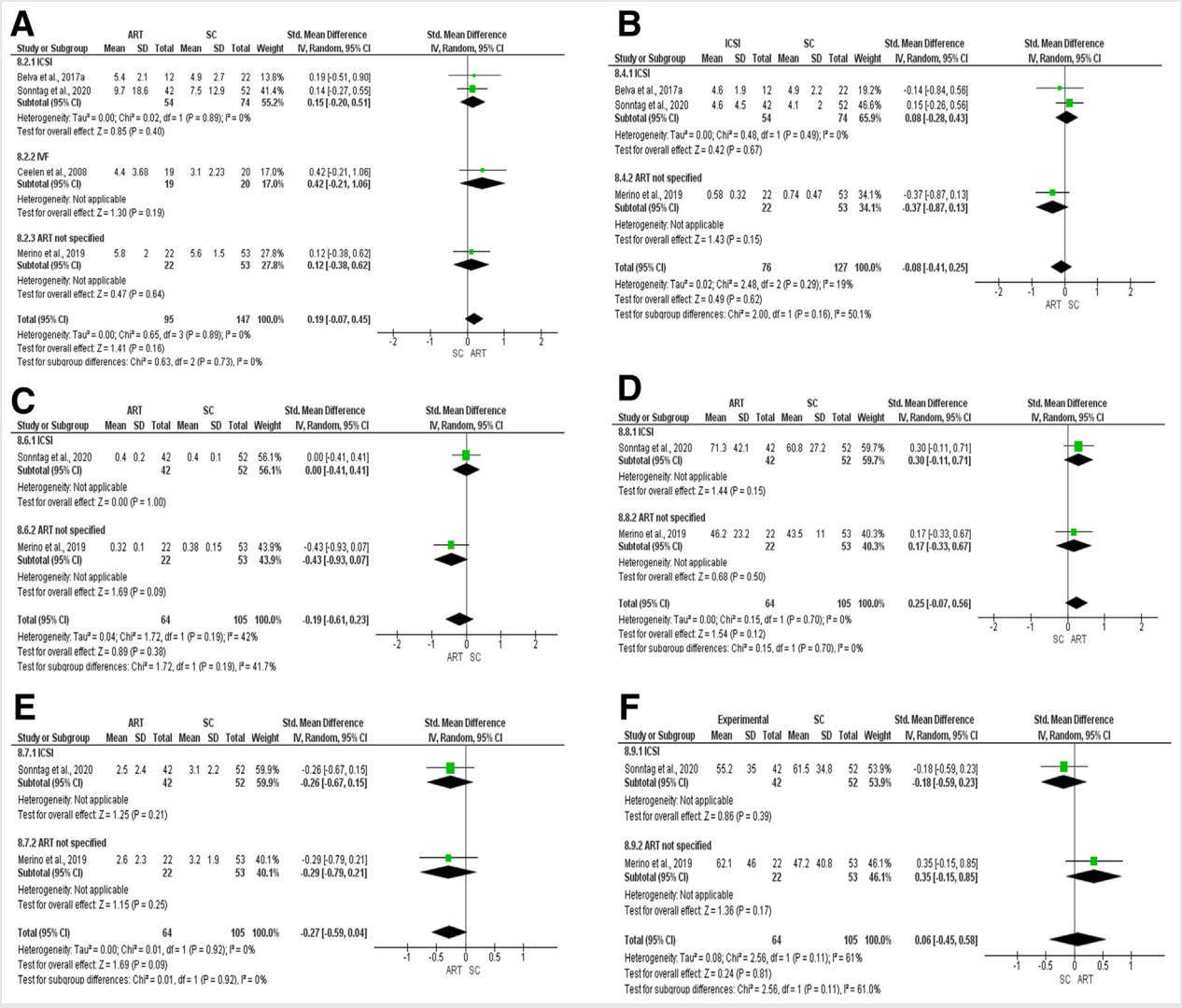
FIGURE 9



(**A**) Forest plot showing standardized mean difference (SMD) in luteinizing hormone (LH) levels between assisted reproductive technique (ART)-conceived and spontaneously conceived (SC) girls. (**B**) Forest plot showing SMD in dehydroepiandrosterone sulfate (DHEAs) levels between ART-conceived and SC girls.

Crafa. ART and offspring gonadal function. Fertil Steril Rev 2023.

FIGURE 10



(**A**) Forest plot showing standardized mean difference (SMD) in follicle-stimulating hormone (FSH) levels between assisted reproductive technique (ART)-conceived and spontaneously conceived (SC) girls. (**B**) Forest plot showing SMD in antiMüllerian hormone (AMH) levels between ART-conceived and SC girls. (**C**) Forest plot showing SMD in total testosterone (TT) levels between ART-conceived and SC girls. (**D**) Forest plot showing SMD in sex hormone-binding globulin (SHBG) levels between ART-conceived and SC girls. (**E**) Forest plot showing SMD in free androgen index (FAI) between ART-conceived and SC girls. (**F**) Forest plot showing SMD in inhibin B levels between ART-conceived and SC girls.

Crafa. ART and offspring gonadal function. Fertil Steril Rev 2023.

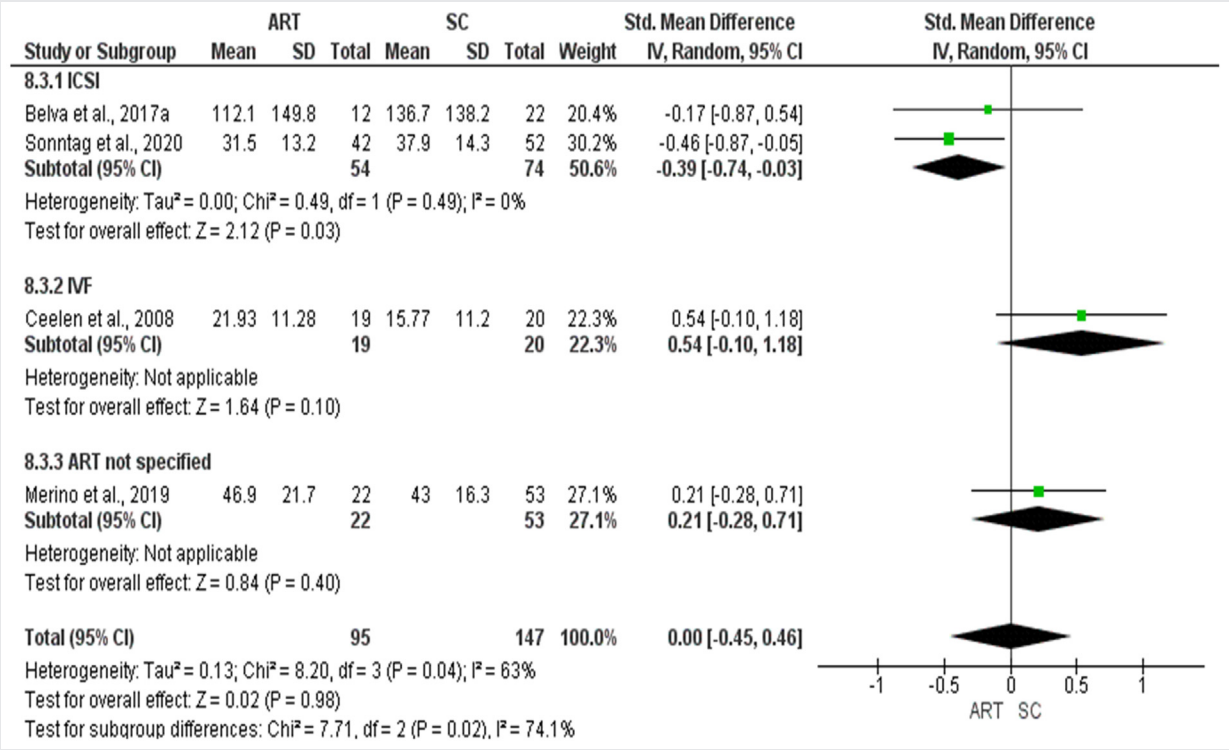
0.011; P=.916; $I^2 = 0$] (Fig. 10E), and inhibin B (SMD = 0.06 (−0.45, 0.58); P=.81; Q-test: Q-value = 2.615; P=.106; I^2 = 61%) (Fig. 10F), independently from the type of ART method used (18, 20, 25, 26). Regarding estrogen levels, only in the subgroup of girls born by ICSI was found a significant reduction compared with the control group (SMD = −0.39 [−0.74, −0.03]; P=.03). The analysis revealed the presence of inter-study heterogeneity (Q-test: Q-value = 8.431; P=.038; I^2 =63%) (Fig. 11) (17, 19, 24, 25). The Egger's regression model and funnel plots reported no risk of bias (intercept = 4.147; 95% CI: −12.86–21,16; P=.20, Supplemental Fig. 5, available online).

Only one single study evaluated androstenedione and found no differences between the groups (25). For the analysis of testosterone, AMH, SHBG, FAI, and inhibin B, no study was sensitive enough to alter the above-reported results (data not shown). The cumulative analysis showed that these results remained nonsignificant after the addition of each study (data not shown).

Sperm Parameters

Only 2 studies evaluated the sperm parameters of youngsters conceived with ART and compared them with SC men (22, 23).

FIGURE 11



Forest plot showing standardized mean difference (SMD) in estradiol levels between assisted reproductive technique (ART)-conceived and spontaneously conceived (SC) girls.

Crafa. ART and offspring gonadal function. Fertil Steril Rev 2023.

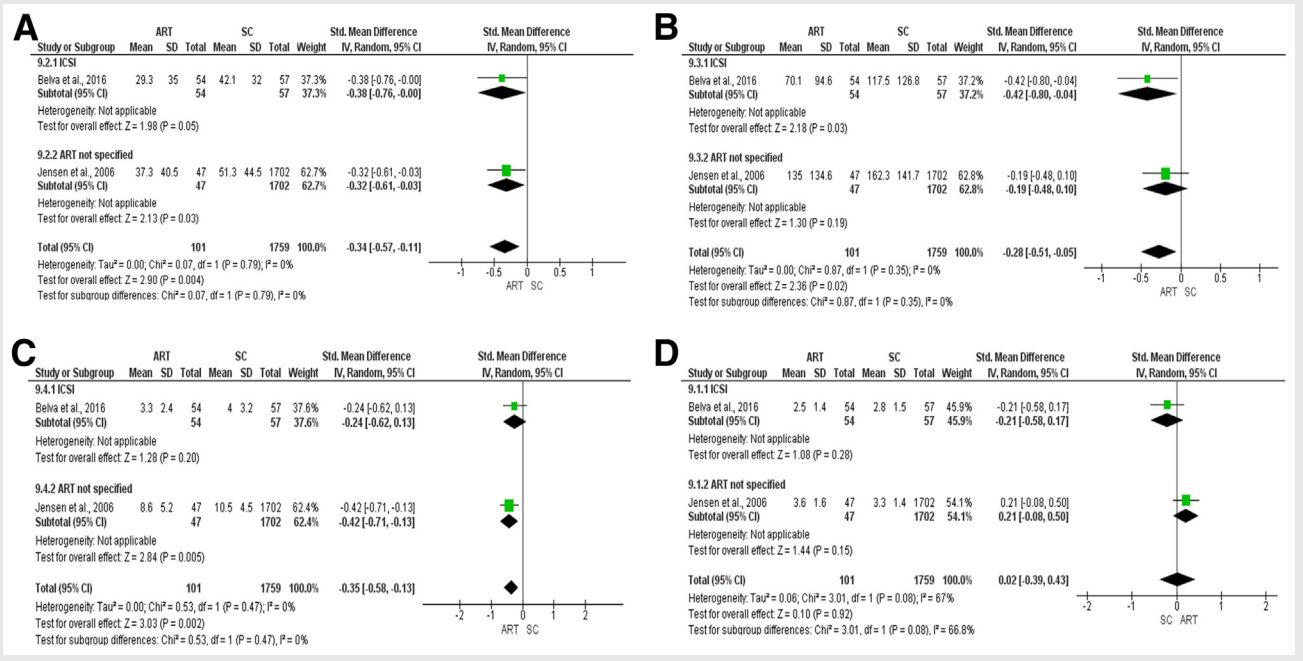
In particular, the quantitative analysis of these 2 studies found that ART-conceived men had significantly reduced sperm concentration (SMD = −0.34 [−0.57, −0.11]; P=.004; Q-test: Q-value = 0,064; P=.80; $I^2 = 0$) (Fig. 12A), total sperm count (SMD = −0.28 [−0.51, −0.05]; P=.02; Q-test: Q-value = 0.894; P=.344; $I^2 = 0$] (Fig. 12B), and sperm normal morphology (SMD = −0.35 [−0.58, −0.13]; P=.002; Q-test: Q-value = 0.519; P =.471; $I^2 = 0$) (Fig. 12C) compared with controls. For the analysis of the total sperm count, the study by Belva et al. (23) was sensitive enough to alter the above-reported results. Its removal resulted in a loss of significance (SMD = −0.193 [−0.487, 0.097]; P=.192). On the other hand, no difference was found for semen volume (SMD = 0.02 [−0.39, 0.43]; P=.92; Q-test: Q-value = 3.035; P=.082; $I^2 = 67\%$) (Fig. 12D). Regarding motility, because the type of motility analyzed is not specified in the study by Jensen et al. (22), it was not possible to perform the quantitative analysis. However, significantly reduced motility was found in this study in men born to mothers who received fertility treatments compared with the control group (22). Instead, in the study by Belva et al. (23), both progressive and total motility were comparable between the 2 groups. However, the total motile sperm count was lower in ICSI-conceived men than in SC men (23).

DISCUSSION

To our knowledge, this is the first meta-analysis that has attempted to evaluate the effects of ART on pubertal development and gonadal function in offspring. To date, it is unclear what mechanisms correlate the use of ART with the overall long-term health of offspring. Even less investigated is the correlation with the reproductive health of the offspring. However, some hypotheses have been proposed that would see a direct and indirect role of ART. Indirect mechanisms include the possibility of direct vertical transmission of male infertility from father to son, as also suggested by some evidence demonstrating the transmission of Y-chromosome microdeletions from infertile fathers to sons (32, 33). Another indirect mechanism appears to be related to birth weight. Indeed, one meta-analytic study observed an increased risk of SGA in children conceived using ART compared with those conceived spontaneously (34). In turn, it was observed that SGA boys and girls on average reached pubertal milestones earlier than those adequate for gestational age (35). Furthermore, SGA is associated with hyperandrogenism and decreased ovarian and uterine volume in women (36) and testicular hypofunction in men (37).

As far as direct mechanisms are concerned, to date, it is still difficult to separate the role of underlying infertility

FIGURE 12



(**A**) Forest plot showing standardized mean difference (SMD) of sperm concentration between assisted reproductive technique (ART)-conceived and spontaneously conceived (SC) boys. (**B**) Forest plot showing SMD of total sperm count between ART-conceived and SC boys. (**C**) Forest plot showing SMD of sperm morphology between ART-conceived and SC boys. (**D**) Forest plot showing SMD of sperm volume between ART-conceived and SC boys.

Crafa. ART and offspring gonadal function. Fertil Steril Rev 2023.

from the procedures performed during ART. It is even more difficult to understand which procedures are associated with the certainty of these alterations (gamete manipulations, culture media, ovarian stimulation, and the use of frozen or fresh embryos) (4). In ICSI, a role may also be because of the invasiveness of the technique itself. In fact, the sperm injection technique could be associated with impaired decondensation of the sperm nucleus, which in turn is associated with embryonic aneuploidies. Furthermore, accidental disruption of the oocyte meiotic spindle, with possible abnormalities in chromosomal segregation or alteration of oocyte pH because of prolonged oocyte handling outside, in turn, associated with an increase in stress-induced aneuploidy rates, could also play a role (38). Among the most studied interfering mechanisms is that of ovarian stimulation. In particular, animal studies have shown that male mice born from mothers who have undergone controlled ovarian hyperstimulation have reduced methylation of the *H19* gene at the level of sperm DNA (39). This finding appears to be of considerable interest because a reduction in the methylation of the *H19* gene has been observed in patients with oligozoospermia, particularly at the level of the CCCTC-binding factor-binding site (40, 41). Another study in mice demonstrated that controlled ovarian hyperstimulation may be associated with delayed pubertal transition and an irregular estrous cycle possibly caused by aberrant growth and maturation of follicles in female offspring (42). All of this evidence suggests that ART may play a role in altering the reproductive function of offspring by various mechanisms. However, these mechanisms still need to be better elucidated.

In our meta-analysis, we found that pubertal development generally proceeds normally and regularly reaches puberty milestones. Consequently, a study of an American cohort of children conceived using IVF found no increase in the rate of precocious or delayed puberty compared with the general population (27). Quantitative analysis showed that in girls alone, there is a higher OR in ICSI-conceived girls with Tanner stage 3 breasts than in SC girls. This suggests the possibility of a slight delay. However, the quantitative analysis is on the basis of only 2 studies and could be influenced by the study by Belva et al. (16) who found breast growth retardation in ICSI-conceived girls compared with girls with SC. However, considering the study's cross-sectional nature, it is impossible to know whether these girls have successfully reached full development.

Hormonally, our meta-analysis showed reduced SHBG levels in ART-conceived boys compared with boys with SC. Sex hormone-binding globulin plays a transport role for sex steroids by regulating circulating concentrations of free hormones and their transport to target tissues (43). However, in our study, the reduction in SHBG levels was not accompanied by significant differences in total and free testosterone levels.

Given the close association of SHBG levels with metabolic syndrome (43, 44) and considering that there is a greater risk of adiposity and insulin resistance in ART-conceived offspring (45), it is possible to hypothesize that the lower levels found in our meta-analysis could be explained by a possible higher prevalence of metabolic syndrome in ART-conceived boys. However, further studies are needed to evaluate this specific aspect.

Quantitative analysis showed the presence of low inhibin B levels only in the subgroup of offspring conceived using ICSI. Inhibin B is produced by Sertoli cells, and its levels reflect the functional state of the seminiferous epithelium (46). Confirming this, quantitative analysis showed a reduction in sperm concentration and sperm count in youngsters conceived using ART. The increased involvement of the ICSI-conceived population could be explained partly by the role of underlying paternal infertility. Indeed, ICSI is used mainly in cases of male infertility, which, in turn, appears to be associated with a negative impact on the long-term health and fertility of the offspring (32).

Regarding females, the quantitative analysis showed higher LH levels in girls conceived using ART than in girls with SC. Elevated LH levels may be associated with hyperandrogenism and the presence of PCOS (47). However, both androgen-level analysis and ultrasound follicle count showed no difference between the 2 groups. Furthermore, the quantitative analysis showed no differences in the risk of developing PCOS in ART-conceived girls compared with SC ones. Finally, the quantitative analysis showed lower estrogen levels only in the subgroup of girls conceived using ICSI compared with girls with SC, suggesting a possible alteration of ovarian function in this group of offspring. However, when interpreting hormonal data in women, it should be taken into account that only Merino et al. (26) specified that hormone tests were performed in the early follicular phase, whereas in the other studies, these data are not presented and could represent a confounding factor.

Several limitations need to be considered when interpreting the data from this meta-analysis. First, because most of the studies are cross-sectional, it is not possible to establish a direct cause-and-effect relationship between ART and the gonadal function of the offspring. Second, the studies are quite heterogeneous from each other, both in the methods of measuring the outcome and in terms of the age of the offspring included. Third, in some studies, there is no exact information on the type of ART method used or the approach used to treat infertility (19, 22, 26). Fourth, the results obtained cannot be attributed with certainty to ART methods per se or underlying parental infertility. Only half of the studies considered the possible interference of parental infertility, with conflicting results (17–19, 21, 23, 28, 30, 31). In particular, 2 studies found no differences in inhibin B levels among children conceived using ICSI on the basis of the severity of paternal infertility (21, 31). Another study reached the same conclusion on salivary testosterone levels (30). In contrast, in the study by Mau Kai et al. (28) on 3-month-old infants, a reduction in testosterone levels was observed only in infants conceived with ICSI from paternal infertility and not in those conceived with IVF from maternal infertility, suggesting a direct role in paternal infertility. Two other studies found no interference in parental infertility in the pubertal development of the offspring and in reaching the pubertal milestones (17, 19). The study by Ceelen et al. (18) showed higher levels of LH and DHEAs in girls born to infertile mothers who underwent IVF compared with those born to infertile mothers who did not receive ovarian stimulation, suggesting a direct role of IVF in determining hormonal alteration. However, even in this case, it must be considered that the analyses were not performed in the follicular phase of the cycle. Finally, only a slight correlation was observed between the total sperm counts of boys conceived using ICSI and those of their fathers, suggesting that the role associated with ICSI in reducing sperm parameters may only be minimally dependent on paternal infertility and could be because of the technique per se (23). This requires studies that directly compare the different techniques with each other. Furthermore, in the context of patients undergoing ICSI, a comparison with children conceived using ICSI using donor sperm could be useful in excluding the role of paternal infertility. Fourth, most of the included studies (17, 21, 23, 25, 29–31) were performed by the same investigators; therefore, there may be some overlap in the population studied. However, data from studies published by the same investigators are never present at the same time in the quantitative analysis, mitigating this limitation. Finally, among the limitations of this meta-analysis, it must be considered that the quantitative analysis is on the basis of a few studies. For this reason, further studies are needed to draw more confident conclusions.

CONCLUSION

This meta-analysis suggests little interference of ART on pubertal development and the achievement of pubertal milestones. Furthermore, Leydig cell function appears to be preserved in the offspring conceived using ART, whereas further studies with correct hormone measurements taken in the early follicular phase are needed to assess ovarian function. On the other hand, scant evidence seems to suggest the interference of ART with testicular germinal function, as shown by the reduction of inhibin B levels and the alteration of sperm parameters. However, further prospective studies are needed to confirm these data and clarify the possible role of the different ART methods and the influence of underlying parental infertility.

REFERENCES

1. Crafa A, Calogero AE, Cannarella R, Mongioì LM, Condorelli RA, Greco EA, et al. The burden of hormonal disorders: a worldwide overview with a particular look in Italy. Front Endocrinol (Lausanne) 2021;12:694325.
2. De Geyter C. Assisted reproductive technology: Impact on society and need for surveillance. Best Pract Res Clin Endocrinol Metab 2019;33:3–8.

3. Barker DJ. The origins of the developmental origins theory. J Intern Med 2007;261:412–7.
4. Chen M, Heilbronn LK. The health outcomes of human offspring conceived by assisted reproductive technologies (ART). J Dev Orig Health Dis 2017;8: 388–402.
5. Cannarella R, Crafa A, Mongioì LM, Leggio L, Iraci N, La Vignera S, et al. DNA methylation in offspring conceived after assisted reproductive techniques: a systematic review and meta-analysis. J Clin Med 2022;11:5056.
6. Cortessis VK, Azadian M, Buxbaum J, Sanogo F, Song AY, Sriprasert I, et al. Comprehensive meta-analysis reveals association between multiple imprinting disorders and conception by assisted reproductive technology. J Assist Reprod Genet 2018;35:943–52.
7. Lazaraviciute G, Kauser M, Bhattacharya S, Haggarty P, Bhattacharya S. A systematic review and meta-analysis of DNA methylation levels and imprinting disorders in children conceived by IVF/ICSI compared with children conceived spontaneously. Hum Reprod Update 2014;20:840–52.
8. Zhang Z, Liu X, Wei C, Luo J, Shi Y, Lin T, et al. Assisted reproductive technologies and the risk of congenital urogenital tract malformations: a systematic review and meta-analysis. J Pediatr Urol 2021;17:9–20.
9. Estill MS, Bolnick JM, Waterland RA, Bolnick AD, Diamond MP, Krawetz SA. Assisted reproductive technology alters deoxyribonucleic acid methylation profiles in bloodspots of newborn infants. Fertil Steril 2016;106:629–39.
10. Jaiswal A, Baliu-Souza T, Turner K, Nadiminty N, Rambhatla A, Agarwal A, et al. Sperm centriole assessment identifies male factor infertility in couples with unexplained infertility - a pilot study. Eur J Cell Biol 2022; 101:151243.
11. Klemetti R, Perry B, Henningsen AKA, Spangmose AL, Pinborg A, Opdahl S, et al. Puberty disorders among ART-conceived singletons: a Nordic register study from the CoNARTaS group. Hum Reprod 2022;37:2402–11.
12. Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ 2015;350:g7647.
13. Murray J, Farrington DP, Eisner MP. Drawing conclusions about causes from systematic reviews of risk factors: the Cambridge Quality Checklists. J Exp Criminol 2009;5:1–23.
14. Hozo SP1, Djulbegovic B, Hozo I. Estimating the mean and variance from the median, range, and the size of a sample. BMC Med Res Methodol 2005;20: 13, 5.
15. Duval S, Tweedie R. A nonparametric "trim and fill" method of accounting for publication bias in meta-analysis. J Am Stat Assoc 2000;95:89–98.
16. Zhu JL, Basso O, Obel C, Bech BH, Nohr EA, Shrestha A, et al. Parental infertility and sexual maturation in children. Hum Reprod 2009;24:445–50.
17. Belva F, Roelants M, Painter R, Bonduelle M, Devroey P, De Schepper J. Pubertal development in ICSI children. Hum Reprod 2012;27:1156–61.
18. Ceelen M, van Weissenbruch MM, Vermeiden JP, van Leeuwen FE, Delemarre-van de Waal HA. Pubertal development in children and adolescents born after IVF and spontaneous conception. Hum Reprod 2008;23:2791–8.
19. Ernst A, Lauridsen LLB, Brix N, Arah OA, Olsen J, Olsen LH, et al. Parental time to pregnancy, medically assisted reproduction and pubertal development in boys and girls. Hum Reprod 2019;34:724–32.
20. Sonntag B, Eisemann N, Elsner S, Ludwig AK, Katalinic A, Kixmüller D, et al. Pubertal development and reproductive hormone levels of singleton ICSI offspring in adolescence: results of a prospective controlled study. Hum Reprod 2020;35:968–76.
21. De Schepper J, Belva F, Schiettecatte J, Anckaert E, Tournaye H, Bonduelle M. Testicular growth and tubular function in prepubertal boys conceived by intracytoplasmic sperm injection. Horm Res 2009;71: 359–63.
22. Jensen TK, Jørgensen N, Asklund C, Carlsen E, Holm M, Skakkebaek NE. Fertility treatment and reproductive health of male offspring: a study of 1,925 young men from the general population. Am J Epidemiol 2007; 165:583–90.
23. Belva F, Bonduelle M, Roelants M, Michielsen D, Van Steirteghem A, Verheyen G, et al. Semen quality of young adult ICSI offspring: the first results. Hum Reprod 2016;31:2811–20.
24. Rojas-Marcos PM, David R, Kohn B. Hormonal effects in infants conceived by assisted reproductive technology. Pediatrics 2005;116:190–4.
25. Belva F, Roelants M, Vloeberghs V, Schiettecatte J, Evenepoel J, Bonduelle M, et al. Serum reproductive hormone levels and ultrasound findings in female offspring after intracytoplasmic sperm injection: first results. Fertil Steril 2017a;107:934–9.
26. Merino PM, Lopez P, Salinas A, Pastene C, Muñoz A, Cassorla F, et al. Ovarian function in adolescents conceived using assisted reproductive technologies. J Pediatr Adolesc Gynecol 2019;32:117–21.
27. Beydoun HA, Sicignano N, Beydoun MA, Bocca S, Stadtmauer L, Oehninger S. Pubertal development of the first cohort of young adults conceived by in vitro fertilization in the United States. Fertil Steril 2011;95:528–33.
28. Mau Kai C, Main KM, Andersen AN, Loft A, Skakkebaek NE, Juul A. Reduced serum testosterone levels in infant boys conceived by intracytoplasmic sperm injection. J Clin Endocrinol Metab 2007;92:2598–603.
29. Belva F, Roelants M, De Schepper J, Van Steirteghem A, Tournaye H, Bonduelle M. Reproductive hormones of ICSI-conceived young adult men: the first results. Hum Reprod 2017b;32:439–46.
30. Belva F, Bonduelle M, Schiettecatte J, Tournaye H, Painter RC, Devroey P, et al. Salivary testosterone concentrations in pubertal ICSI boys compared with spontaneously conceived boys. Hum Reprod 2011;26: 438–41.
31. Belva F, Bonduelle M, Painter RC, Schiettecatte J, Devroey P, De Schepper J. Serum inhibin B concentrations in pubertal boys conceived by ICSI: first results. Hum Reprod 2010;25:2811–4.
32. Rumbold AR, Sevoyan A, Oswald TK, Fernandez RC, Davies MJ, Moore VM. Impact of male factor infertility on offspring health and development. Fertil Steril 2019;111:1047–53.
33. Calogero AE, Garofalo MR, Barone N, Longo GA, De Palma A, Fichera M, et al. Spontaneous transmission from a father to his son of a Y chromosome microdeletion involving the deleted in azoospermia (DAZ) gene. J Endocrinol Invest 2002;25:631–4.
34. Qin J, Liu X, Sheng X, Wang H, Gao S. Assisted reproductive technology and the risk of pregnancy-related complications and adverse pregnancy outcomes in singleton pregnancies: a meta-analysis of cohort studies. Fertil Steril 2016;105:73–85.
35. Hvidt JJ, Brix N, Ernst A, Lauridsen LLB, Ramlau-Hansen CH. Size at birth, infant growth, and age at pubertal development in boys and girls. Clin Epidemiol 2019;11:873–83.
36. Petraitiene I, Valuniene M, Jariene K, Seibokaite A, Albertsson-Wikland K, Verkauskiene R. Sex hormones, gonad size, and metabolic profile in adolescent girls born small for gestational age with catch-up growth. J Pediatr Adolesc Gynecol 2020;33:125–32.
37. Cicognani A, Alessandroni R, Pasini A, Pirazzoli P, Cassio A, Barbieri E, et al. Low birth weight for gestational age and subsequent male gonadal function. J Pediatr 2002;141:376–9.
38. Esteves SC, Roque M, Bedoschi G, Haahr T, Humaidan P. Intracytoplasmic sperm injection for male infertility and consequences for offspring. Nat Rev Urol 2018;15:535–62.
39. Stouder C, Deutsch S, Paoloni-Giacobino A. Superovulation in mice alters the methylation pattern of imprinted genes in the sperm of the offspring. Reprod Toxicol 2009;28:536–41.
40. Li B, Li JB, Xiao XF, Ma YF, Wang J, Liang XX, et al. Altered DNA methylation patterns of the H19 differentially methylated region and the DAZL gene promoter are associated with defective human sperm. PloS one 2013;8: e71215.
41. Kobayashi H, Sato A, Otsu E, Hiura H, Tomatsu C, Utsunomiya T, et al. Aberrant DNA methylation of imprinted loci in sperm from oligospermic patients. Hum Mol Genet 2007;16:2542–51.
42. Ding J, Tan X, Song K, Ma W, Xiao J, Zhang M. Effect of controlled ovarian hyperstimulation on puberty and estrus in mice offspring. Reproduction 2017;154:433–44.
43. Wallace IR, McKinley MC, Bell PM, Hunter SJ. Sex hormone binding globulin and insulin resistance. Clin Endocrinol (Oxf) 2013;78:321–9.
44. Brand JS, Rovers MM, Yeap BB, Schneider HJ, Tuomainen TP, Haring R, et al. Testosterone, sex hormone-binding globulin and the metabolic syndrome in

men: an individual participant data meta-analysis of observational studies. PloS one 2014;9:e100409.
45. Heber MF, Ptak GE. The effects of assisted reproduction technologies on metabolic health and disease. Biol Reprod 2021;104:734–44.
46. Meachem SJ, Nieschlag E, Simoni M. Inhibin B in male reproduction: pathophysiology and clinical relevance. Eur J Endocrinol 2001;145:561–71.
47. Coutinho EA, Kauffman AS. The role of the brain in the pathogenesis and physiology of polycystic ovary syndrome (PCOS). Med Sci (Basel) 2019;7:84.
48. Jørgensen N, Auger J, Giwercman A, Irvine DS, Jensen TK, Jouannet P, et al. Semen analysis performed by different laboratory teams: an intervariation study. Int J Androl 1997;20:201–8.