Effect of semen collection location on semen parameters and fertility outcomes—implications for practice in the COVID-19 era: A systematic review and meta-analysis of randomized and observational studies

Pakapong Kerdtawee, Lingling Salang, Jen Sothornwit

PII: S0002-9378(22)00726-8
DOI: https://doi.org/10.1016/j.ajog.2022.09.009
Reference: YMOB 14729


Received Date: 21 April 2022
Revised Date: 20 August 2022
Accepted Date: 2 September 2022


This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2022 Elsevier Inc. All rights reserved.
Effect of semen collection location on semen parameters and fertility outcomes–implications for practice in the COVID-19 era: A systematic review and meta-analysis of randomized and observational studies

Pakapong Kerdtawee, Lingling Salang*, Jen Sothornwit

Department of Obstetrics and Gynecology, Faculty of Medicine, Khon Kaen University, Khon Kaen, 40002, Thailand

*Correspondence: Lingling Salang

Department of Obstetrics and Gynecology, Faculty of Medicine, Khon Kaen University, Khon Kaen, 40002, Thailand

Tel +6643363030
Fax +6646348395

Email: Slingling@kku.ac.th

The authors reported no conflict of interest.
Condensation

At-home semen collection had no adverse effect on semen parameters or fertility outcomes compared to in-clinic collection, but higher quality evidence is needed.

Short Title

Semen collection location on parameters and outcomes.

AJOG at a Glance

Why was this review conducted?

This review was conducted to assess the effects of semen collection location on semen parameters and fertility outcomes.

Key findings

This meta-analysis showed that at-home semen collection had no adverse effects on semen parameters or fertility outcomes compared to at-clinic semen collection.

What does this add to what is known?

This is the first systematic review and meta-analysis conducted with all currently available data and large samples comparing the effect of at-home vs. at-clinic semen collection on semen parameters and clinical outcomes, which showed comparable results. High-quality evidence from Randomized control trials (RCTs) is needed to strengthen the evidence for future practice.
Abstract

Objective: During the COVID-19 era, semen collection at infertility centers might increase the risk for spreading SARS-CoV-2. Therefore, seminal fluid collection at home is an alternative method for preventing this spread. However, there is no conclusion about the effect of home versus clinic semen collection on semen parameters and assisted reproductive technology (ART) outcomes. This systematic review and meta-analysis aimed to assess the effect of semen collection on semen parameters and fertility outcomes.

Data Sources: A literature search was conducted using the major electronic databases MEDLINE via Ovid, EMBASE, Scopus, CINAHL, OpenGrey, and CENTRAL from their inception to September 2021. ClinicalTrials.gov was searched to identify the ongoing registered clinical trials.

Study Eligibility Criteria: We included all human randomized controlled trials and observational studies that investigated the effect of at-home semen collection versus in-clinic semen collection on semen parameters and fertility outcomes.

Methods: We pooled the mean difference and risk ratio using Review Manager software version 5.4.1. The GRADE approach was applied to assess the quality of evidence.

Results: Seven studies (3018 semen samples) were included. Overall, at-home semen collection results made little to no difference in semen volume (mean difference [MD] 0.37; 95% CI -0.10 to 0.85; low-quality evidence), sperm count (MD -6.02; 95% CI -27.26 to 15.22; very low-quality evidence), and sperm motility (MD 0.76; 95%CI -4.39 to 5.92; very low-quality evidence) compared to in-clinic semen collection. There was no difference in fertilization rate (risk ratio
[RR] 1.00; 95% CI 0.97 to 1.03; very low-quality evidence) and pregnancy rate in in vitro fertilization (RR 1.04; 95% CI 0.86 to 1.25; very low-quality evidence).

**Conclusion**: At-home semen collection had no adverse effects on semen parameters or fertility outcomes compared to in-clinic collection; however, higher-quality evidence is needed.

**Key Words**: semen analysis, collection location, semen parameters, fertility outcomes
INTRODUCTION

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (COVID-19) remains a serious threat to public health. Its rapid transmission has led to worldwide spread of the disease over the last two years. The World Health Organization (WHO) declared COVID-19 a global pandemic with widespread implications for the healthcare system (1). Inevitably, infertility management is affected by avoidance of human physical contact and decreased clinic visits, as cited by the American Society for Reproductive Medicine (ASRM), European Society of Human Reproduction and Embryology (ESHRE), and International Federation of Fertility Societies (IFFS) (2-4).

There is inconclusive evidence regarding the effect of COVID-19 infection on the male reproductive organs in a short and long term. Focusing on semen parameters, Guo et al. reported no impact of COVID infection on sperm concentration, motility, and morphology (5), while other studies showed a trend in negative effect on some seminal parameters (6-8).

Generally, males are assigned to an infertility center for semen collection, which might increase the risk of spreading SARS-CoV-2. Collecting semen at home is an alternative strategy to prevent the spread of SARS-CoV-2. The WHO manual recommends that, in exceptional cases, semen can be collected at home within an hour of analysis (9-12). Gao et al. reported on the benefits of home semen collection, stating that males who collected semen at home were more relaxed and reached orgasm more easily than males who collected semen in clinics (13). There is no conclusion about the effect of home versus clinic semen collection on semen parameters and assisted reproductive technology (ART) outcomes.
Various studies have reported the effects of semen collection location on semen parameters and infertility treatment outcomes. For example, several studies reported that home semen collection significantly increased the mean time to semen processing compared to in-clinic collection; however, semen parameters and pregnancy rates in intrauterine insemination (IUI) and in vitro fertilization (IVF) outcomes were not significantly different between groups(13-15). In contrast, Yavas et al. reported differences in IUI cycle outcomes, viz., pregnancy rate was significantly higher with in-clinic semen collection than with at-home hMG stimulation (44 vs. 18%; P-value = 0.03); however, the difference in clomiphene citrate stimulation was not statistically significant(16).

During the pandemic, at-home semen collection could be a satisfactory alternative, whether for work-up, IUI, or ART. The current review was undertaken to assess the effects of semen collection location on (a) seminal parameters (semen volume, sperm concentration, total sperm count, sperm motility, total motile sperm count, progressive motile sperm, and normal sperm morphology); and (b) fertility outcomes (fertilization rates, usable blastocyst, pregnancy, miscarriage, and live birth rates).

MATERIALS AND METHODS

Eligibility Criteria, Information Sources, and Search Strategy

A systematic literature review was conducted following the Cochrane Handbook for Systematic Reviews of Interventions(17) and was reported according to the Meta-analysis of Observational Studies in Epidemiology (MOOSE)(18). The review protocol was registered with PROSPERO (CRD42021268246).
To identify potentially eligible studies, a systematic literature search was conducted by authors using the major electronic databases from their inception to September 2021, including MEDLINE via Ovid, EMBASE, Scopus, CINAHL, and the Cochrane Central Register of Controlled Trials Database (CENTRAL) (Supplemental Appendices). The reference lists of articles were checked, and authors of the trials were contacted to obtain additional data if necessary. ClinicalTrial.gov and the World Health Organization Internal Clinical Trials Registry Platform (http://www.who.int/ictrp) were searched for unpublished, planned, and ongoing trial reports. Open Grey (http://www.opengrey.eu) was used to search for grey literature.

The PICO-S elements of this review were population, any man, intervention, at-home semen collection, comparison, and in-clinic semen collection, and outcomes of interest were the location of semen collection, semen parameters (semen volume, sperm concentration, total sperm count, sperm motility, total motile sperm count, progressive motile sperm, and normal sperm morphology), fertility outcomes (fertilization rates, usable blastocyst rates, pregnancy rates, miscarriage rates, and live birth rates), and study design. We included randomized controlled trials (RCTs) and observational studies published as full-reports or conference abstracts regardless of the language of publication, publication status, year of publication, or sample size.

**Study Selection and Data Extraction**

The titles and abstracts of the studies retrieved by electronic search were screened independently by two authors. Titles and abstracts that did not meet the inclusion criteria were excluded. Two authors retrieved and independently reviewed the full texts for potentially eligible studies and
extracted relevant data. Any disagreements on relevance were resolved through consensus with a third person.

**Risk of Bias Assessment**

Two tools were used to evaluate the risk of bias: (a) the revised Cochrane risk of bias tool for randomized trials (RoB 2)(19) for RCTs; and (b) the risk of bias in non-randomized studies of interventions (ROBINS-I)(20). For observational studies, two authors independently assessed the risk of bias. Any disagreements were resolved by consensus with a third reviewer.

**Data Synthesis**

The data extracted from the review were pooled and analyzed using Review Manager software version 5.4.1(21). Means with standard deviations (SD) or converted means and SDs in cases where studies provided data as medians, ranges, and interquartile ranges using a standard formula were used in the meta-analysis(22). Outcomes were measured as mean difference with a 95% confidence interval or risk ratio with a P-value of <0.05 considered statistically significant. Heterogeneity of the studies was assessed using $I^2$, which defined values between 75% and 100%, and a chi-square test with a significance level of 0.10 were considered as heterogeneous(23). When heterogeneity was high, a random effects model was used. Subgroup or sensitivity analyses were planned for a specific population or study type to assess the robustness of the results.

**Quality of Evidence**

The certainty of the body of evidence was assessed using the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) approach. The GRADE approach covered the following five domains: 1) risk of
bias in the included studies; 2) inconsistency between studies’ 3) imprecision in
the effect estimate; 4) indirectness of evidence; and 5) publication bias. The
GRADE approach rates the overall certainty of evidence as high, moderate,
low, or very low-quality (24).

RESULTS

Study Selection

Figure 1 presents the PRISMA flowchart for the study selection. A broad
search yielded 1,229 reports from electronic database searches and other
sources. After removing duplicate, 903 reports remained for screening, and
892 that did not meet the inclusion criteria were excluded. After reviewing the
full texts of 11 reports that potentially met the review inclusion criteria, one
was excluded because the full text was unavailable. Finally, 10 reports from 7
studies, involving 3,018 samples were included in the qualitative synthesis.

Characteristics of Included Studies

A total of seven studies(13-16, 25-27) were included for this review
ranging between 2004-2021. There were six observational studies(14-16, 25-
27), most of which were retrospective, and one was a RCT(13). Three studies
reported only semen parameters(13, 14, 25), one reported only fertility
outcomes(16), and three reported both semen parameters and fertility
outcomes(15, 26, 27). The detailed characteristics of the studies are presented
in Table 1.

Risk of Bias in the Included Studies

Figures 2 and 3 present a summary of the risk of bias in each included
study. The risk of bias for six studies(14-16, 25-27) was assessed using the
ROBIN-I tool because they were observational studies. The risk of bias for the
RCT(13) was assessed using the Cochrane RoB2 tool.

With respect to observational studies, the four included studies(15, 16,
26, 27) were assessed to have a moderate risk of bias due to confounding
factors, such as age, days of abstinence, sperm quality, and laboratory
methods. The confounding factors were appropriately controlled so that they
did not pose a severe residual effect. In two of the six included studies(14, 25),
bias was deemed serious because there was no mention of how confounders
could be controlled.

Vis-à-vis domain selection bias, all studies were judged to have a
moderate risk of bias because they were observational. In studies in which
participants could choose which group they preferred (i.e., without
randomization), the outcome could have been skewed(14-16, 25-27).

All studies were judged to have a low risk of bias vis-à-vis intervention
classification, deviations from intended interventions, missing data,
measurement of outcomes, and selection of reported result domain(14-16, 25-
27).

In the case of RCTs, only one study was evaluated using the Cochrane
RoB2 tool; however, all domains of the risk of bias (e.g., selection,
performance, detection, attrition, and reporting) were judged to have a low
risk of bias.
Synthesis of Results

Primary Outcome

Six studies (five observational studies and one RCT) reported semen parameters (13-15, 25-27). Three studies compared overall at-home collected semen parameters with at-clinic collected parameters (25-27). Only one study reported individual sample outcomes between the at-home and in-clinic parameters from the same participants (14), whereas two studies (13, 15) reported overall and individual semen parameters. All studies reported basic semen parameters, either as means with SDs or medians with interquartile ranges, depending on data distribution. After requesting for outcome data as means ± SD, only one study (25) was converted to means using Wan’s formula (22).

Six studies (13-15, 25-27)—including 3018 semen samples, were included in the meta-analysis of semen parameters (semen volume, sperm concentration, total sperm count, sperm motility, total motile sperm count, progressive motile sperm, and normal sperm morphology). Comparisons between at-home and in-clinic semen collections are presented as a forest plot showing the mean difference (Figure 4).

Semen Volume

The semen volume did not differ significantly when it was collected at home or in the clinic. Subgroup analysis indicated no significant difference in semen volume between the two comparison groups when stratified by overall (mean difference [MD] 0.37; 95% CI -0.10 to 0.85; 4 studies; 1806 samples) (13, 15, 25, 27) and individual comparisons (MD -0.08; 95% CI -0.38 to 0.22; 3 studies; 628 samples) (13-15) (Figure 4A).
Sperm Concentration

There was no significant difference between sperm concentration parameters collected at home and in the clinic, or between overall samples (MD, 5.46; 95% CI -1.23 to 12.15; 5 studies; 2439 samples)(13, 15, 25-27) and individual samples (MD0.82; 95% CI -7.09 to 8.73; 3 studies; 628 samples)(13-15) (Figure 4B).

Total Sperm Count

Overall, semen collection at home did not reduce the total sperm count. Additionally, the subgroup analysis indicated no significant benefit in collecting semen at home vis-à-vis increasing semen concentration overall (MD 14.16; 95% CI -2.88 to 31.20; 3 studies; 1562 samples)(13, 25, 27) or individual comparisons (MD -7.34; 95% CI –33.43 to 18.76; 2 studies; 544 samples)(13, 14) (Figure 4C).

Sperm Motility

The evidence was ‘very uncertain’ on sperm motility regarding the effect of collecting semen at home vs. in clinic. Subgroup analysis indicated no significant benefit in semen collection at home vis-à-vis increasing sperm motility whether overall (MD 0.76; 95% CI -4.39 to 5.92; 3 studies; 1958 samples)(15, 26, 27) or individual comparisons (MD -0.55; 95% CI -3.67 to 2.58; 2 studies; 424 samples)(14, 15) (Figure4D).

Total Motile Sperm Count

In the respective overall subgroup and individual analyses, collecting semen at home vs. in clinic did not appear to increase the total motile sperm
Progressive Motility

There was no significant difference in progressive motility between home and clinic collected semen groups, regardless of the overall sample (MD 1.17; 95% CI -2.42 to 4.76; 3 studies; 1114 samples)(13, 25, 26) or individual samples (MD 1.02; 95% CI -2.94 to 4.99; 1 study; 204 samples)(13) (Figure 4E).

Normal Morphology

Evidence suggests that collecting semen at home versus in the clinic does not change the normal sperm morphology. Subgroup analysis indicated no significant benefit to either location, whether overall samples (MD -0.10; 95%CI -0.66 to 0.46; 3 studies; 1562 samples)(13, 25, 27) or individual samples (MD 0.14; 95% CI -0.27 to 0.55; 2 studies; 544 samples)(13, 14) (Figure 4F).

Secondary Outcome

The scope of fertility outcomes includes fertilization, usable blastocyst, pregnancy, miscarriage, and live birth rates. Two studies(15, 27) were conducted with IVF/ICSI populations and reported fertilization rates (defined as the number of two pronuclear stage embryos/number of metaphase II oocytes) with a total of 10,109 metaphase II oocytes and pregnancy rates (defined as the number of positive beta-hCG tests/embryo transfers) with a total of 976 transfers. The other two studies(16, 26) were conducted among people who underwent IUI and reported only pregnancy rates (defined as an ultrasonographic finding of fetal cardiac activity/number of IUIs) with a total of 765 IUIs. Thus, a meta-analysis of pregnancy rates was conducted with these specific IVF/ICSI and IUI subgroups. Two studies(15, 27) reported usable
blastocyst rates, but there were differences in the definitions and methods of outcome measurement. Sacha et al. (15) defined the number of day five transferable and freezable quality blastocysts/number of two pronuclear stage embryos, but Stimpfel et al. (27) defined it as the proportion of transferred/cryopreserved embryos (referred to as day 5/6 blastocysts) per number of embryos obtained. Consequently, we could not conduct a meta-analysis for this outcome, and there were no data on miscarriage and live birth rates in any of the included studies; therefore, a meta-analysis could not be performed on these parameters either. Forest plots of the estimated fertility outcomes are shown in Figure 3.

Fertilization Rates

The evidence suggests that collecting semen at home vs. in clinic did not increase fertilization rates in the IVF cycle (RR 1.00; 95% CI 0.97 to 1.03; 2 studies; 6770 events) (15, 27).

Pregnancy Rates

The evidence is very uncertain in relation to the effect of collecting semen at home on the pregnancy rate in the IUI and IVF/ICSI cycles (RR 0.85; 95% CI 0.28 to 2.59; 2 studies (16, 26); 76 events; and RR 1.04; 95% CI 0.86 to 1.25; 2 studies (15, 27); 356 events, respectively).

Quality of the Studies

The quality of evidence for all semen parameters was very low, owing to selection bias, lack of blinding, and imprecision of estimation in the included studies. We rated the quality of evidence for the fertility rate as very low due to a lack of information on participant selection and blinding in the included studies and small sample sizes. We rated the quality of evidence as very low
for the pregnancy rate, owing to a lack of blinding, small sample size, and selection bias.

Comment

Principal Findings

This present review is the first systematic review and meta-analysis to estimate the effect of location on (home versus clinic collected) semen parameters and fertility outcomes—estimated average from 3,018 semen samples from participants between 20-58 years of age. We analyzed and reported estimated outcomes in subgroup comparisons as overall and individualized samples but could not report the summarized outcome as total estimated because individual data were a subset of overall data in some studies (13, 15).

The results of the meta-analysis for each subgroup showed that the semen volume, sperm concentration, total sperm count, sperm motility, total motile sperm count, progressive motile sperm, and normal sperm morphology were not negatively affected by the location of semen collection. There are concerns that transport time of semen collected at home to the laboratory may affect semen or sperm quality because of (a) increased exposure of spermatozoa to seminal plasma; or (b) changes in temperature during transport. However, no negative impact on at-home collected semen was found for any semen parameters, regardless of whether the ICSI/IVF or IUI cycles were used. These results might be explained by the semi-controlled effect of some confounders, such as length of abstinence (normally 2-7 days)(25), but can range between 1-30 days (median four days). According to the WHO guidelines, most studies report the time to transfer samples to the
laboratory on time, but three studies (14, 15, 27) reported that the time to transport was between 1.5-2 h with no negative impact on semen quality.

Other potential confounding effects on outcomes included baseline underlying medical conditions, endocrine system health, psychological status, medications, or substances used, and methods used to collect semen. Additionally, the techniques, instruments, and procedures used to analyze each semen parameters (13-15, 27), such as conventional manual procedures or computer-aided sperm analysis (CASA), could be confounders.

During this present COVID-19 pandemic, there were more confounding variables from COVID-19 related factors, such as history of viral infection, which may have effects on the male reproductive system from proposed hypothesis that viruses directly damage the target cell, or inflammatory response by cytokines, or testicular damage from fever (28). There was a report of low level of testosterone after recovery from COVID-19, especially in patients with history of severe symptoms (29). Additionally, in a study by Gonzalez et al. (30), the status of COVID-19 mRNA vaccination may have an effect on semen parameters, which showed significant increase in sperm parameters. However, small and larger samples are needed for effective conclusions. None of the included studies provided information about this potential confounder.

Only two included studies (15, 27) reported the adjusted outcomes and added male age, while one study reported the number of days of abstinence and added female age, fertilization method, and the number of oocytes retrieved, or embryos transferred for fertility outcomes. Other studies did not mention the potential confounding factors mentioned above, which might
affect the certainty of the estimated outcome and be a limitation of the review.

Sperm parameters and clinical outcomes can be proxied for infertility outcomes. The fertilization rate was comparable between the two groups with no heterogeneity in the two studies, and the pregnancy rates, even in the subgroups of those who underwent ICSI/IVF or IUI, were not significantly different between the at-home and in-clinic semen collections. Another point of concern is the effect of conventional IVF or ICSI on fertilization and pregnancy rates. Only one of the included studies by Stimpfel et al. (27) showed results regarding fertility outcomes (only pregnancy rate) in IVF and ICSI subgroups. After we performed a meta-analysis, there was no difference in pregnancy rate between the two sites of semen collection when only conventional IVF cycles or ICSI cycles were included. These results may be due to the non-significant differences in the semen parameters used in the ART procedure. For the estimated outcome of pregnancy rates in the IUI subgroup, there was high heterogeneity; therefore, we explored this and assumed that the difference in outcome might be due to differences in methodology and confounding factors, such as baseline female factor, protocol, and the procedure used, which were not adjusted in some studies.

The reason these confounders were considered significant is that they may have contributed to the considerable heterogeneity, in addition to the differences in design and methodology for each study. The overall quality of evidence was another factor determined by this meta-analysis. To this end, we used the GRADE system for cohort-type studies according to the study design itself. Considerable heterogeneity was found in semen volume and motility outcomes. Heterogeneity among studies often occurred within a subgroup of
the overall sample, while within-person comparisons had no heterogeneity. This may be due to differences in methodology and baseline characteristics, leading to interpersonal-variation effects. The implications of the estimated outcomes from the current review should be generalized with caution because of the limitations of the review.

**Strengths and Limitations**

This is the first systematic review and meta-analysis conducted with all currently available data and large samples comparing the effect of at-home vs. in-clinic semen collection on semen parameters and clinical outcomes (i.e., pregnancy rates). A systematic review was conducted following the Cochrane and MOOSE guidelines. Additionally, subgroup analysis was used to reduce bias and the effects of heterogeneity among studies and to estimate robustness.

A limitation of this review is that most of the included studies were observational, with only one RCT. A low-quality study design eroded the quality of evidence from RCT, resulting in greater heterogeneity among studies. Furthermore, most studies did not adjust for potential confounders, which may have affected the outcomes of the original studies. A lack of intermediate and long-term clinical outcomes exists, for example, miscarriage and live birth rates; therefore, more high-quality studies, such as RCTs are needed to strengthen the evidence for future practice.

**Comparison with Existing Literature**

This review is the first systematic review and meta-analysis to estimate the outcome of location of semen collection and whether it affects semen
parameters and fertility outcomes. Thus, there are no existing reviews available for comparison with this review.

**Conclusions and Implications**

Evidence indicates that collecting semen at home did not result in any significant difference in semen parameters, fertility rates, or pregnancy rates. Further studies should include more high-quality RCTs. The outcomes of this evidence-based meta-analysis support at-home semen collection as a qualitatively acceptable option, particularly during the COVID-19 pandemic, which may play a role in future routine ART services.

**Author contributions**

All authors have read and prepared the manuscript. We thank Mr. Bryan Roderick Hamman under the aegis of the Khon Kaen Publication Clinic for assistance with the English language presentation of the manuscript.

**Declaration of Competing Interest**

The authors declare that they have no known competing financial interests or personal relationships that could have influenced the work reported in this study.
References


13. Gao J, Duan YG, Yi X, Yeung WSB, Ng EHY. A randomised trial comparing conventional semen parameters, sperm DNA fragmentation levels and satisfaction levels between semen collection at home and at the clinic. Andrologia. 2020;52(7):e13628.


16. Yavas Y, Selub MR. Intrauterine insemination (IUI) pregnancy outcome is enhanced by shorter intervals from semen collection to sperm wash, from sperm wash to IUI time, and from semen collection to IUI time. Fertil Steril. 2004;82(6):1638-47.


Table 1. Characteristics of included studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Year</th>
<th>Country</th>
<th>Study design</th>
<th>Home (n)</th>
<th>Clinic (n)</th>
<th>Inclusion criteria</th>
<th>Exclusion criteria</th>
<th>Study outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elzanaty et al.</td>
<td>2008</td>
<td>Sweden</td>
<td>Cross-sectional study</td>
<td>106</td>
<td>273</td>
<td>Men undergoing infertile assessment</td>
<td>azoospermia</td>
<td>volume, concentration, total count, progressive motility, morphology</td>
</tr>
<tr>
<td>Gao et al.</td>
<td>2020</td>
<td>China</td>
<td>RCT</td>
<td>53</td>
<td>49</td>
<td>Men 18-55 years, infertility for at least 1 year</td>
<td>- presence of dysuria, urinary urgency, and increased frequency of urination - erectile or ejaculatory dysfunction - inability to follow instructions due to impaired cognition</td>
<td>volume, concentration, total count, progressive motility, morphology</td>
</tr>
<tr>
<td>Licht et al.</td>
<td>2008</td>
<td>USA</td>
<td>prospective cohort study</td>
<td>232</td>
<td>215</td>
<td>Men undergoing work up for infertility and for IUI</td>
<td>not mentioned</td>
<td>volume, concentration, total count, motility, total motile count, morphology</td>
</tr>
<tr>
<td>Sacha et al.</td>
<td>2021</td>
<td>USA</td>
<td>retrospective cohort study</td>
<td>125</td>
<td>119</td>
<td>Men undergoing for IVF/ICSI</td>
<td>not mentioned</td>
<td>volume, concentration, motility, total motile count, morphology</td>
</tr>
<tr>
<td>Song et al.</td>
<td>2007</td>
<td>USA</td>
<td>retrospective study</td>
<td>236</td>
<td>397</td>
<td>Men undergoing for IUI</td>
<td>not mentioned</td>
<td>total count, motility, total motile count, progressive motility, pregnancy rate</td>
</tr>
<tr>
<td>Stimpfel et al.</td>
<td>2021</td>
<td>Slovenia</td>
<td>retrospective study</td>
<td>837</td>
<td>244</td>
<td>Men undergoing for IVF/ICSI</td>
<td>not mentioned</td>
<td>volume, concentration, total count, motility, morphology, fertilization rate, blastocyst rate, embryo utilization rate, pregnancy rate</td>
</tr>
<tr>
<td>Yavas et al.</td>
<td>2004</td>
<td>USA</td>
<td>retrospective study</td>
<td>95</td>
<td>37</td>
<td>Men undergoing for IUI</td>
<td>not mentioned</td>
<td>pregnancy rate</td>
</tr>
</tbody>
</table>
Figure 1. PRISMA 2020 flow diagram for systematic review and meta-analysis
**Figure 2.** Risk of bias assessment of RCTs by RoB 2 tool

<table>
<thead>
<tr>
<th>Study</th>
<th>Risk of bias domains</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gao (2019)</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

Domains:
- D1: Bias arising from the randomization process.
- D2: Bias due to deviations from intended intervention.
- D3: Bias due to missing outcome data.
- D4: Bias in measurement of the outcome.
- D5: Bias in selection of the reported result.
**Figure 3.** Risk of bias assessment of observational studies by ROBIN-I tool

<table>
<thead>
<tr>
<th>Study</th>
<th>D1</th>
<th>D2</th>
<th>D3</th>
<th>D4</th>
<th>D5</th>
<th>D6</th>
<th>D7</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elzanaty (2008)</td>
<td>X</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>X</td>
</tr>
<tr>
<td>Licht (2008)</td>
<td>X</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>X</td>
</tr>
<tr>
<td>Sacha (2021)</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Song (2007)</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Stimpfel (2021)</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Yavas (2004)</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

**Domains:**
- D1: Bias due to confounding.
- D2: Bias due to selection of participants.
- D3: Bias in classification of interventions.
- D4: Bias due to deviations from intended interventions.
- D5: Bias due to missing data.
- D6: Bias in measurement of outcomes.
- D7: Bias in selection of the reported result.

**Judgement**
- Red: Serious
- Yellow: Moderate
- Green: Low
Figure 4. Estimates semen parameters

Figure 4A. comparison in semen volume

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Home collected Mean</th>
<th>Home collected SD</th>
<th>Home collected Total</th>
<th>Clinic collected Mean</th>
<th>Clinic collected SD</th>
<th>Clinic collected Total</th>
<th>Mean Difference</th>
<th>Weight</th>
<th>IV, Random, 95% CI</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall samples</td>
<td>6.25 2.779 106</td>
<td>5.1763 273</td>
<td>21.7% 1.25 [0.68, 1.82]</td>
<td>2008</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elzanaty 2008</td>
<td>3.891 1.583 53</td>
<td>3.426 1.397</td>
<td>49 21.5% 0.46 [-0.11, 1.04]</td>
<td>2020</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sacha 2021</td>
<td>2.8 1.6 125</td>
<td>2.8 1.3 119</td>
<td>26.5% 0.00 [-0.37, 0.37]</td>
<td>2021</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simpel 2021</td>
<td>2.6 1.2 837</td>
<td>2.6 1.1 244</td>
<td>30.3% 0.00 [-0.16, 0.16]</td>
<td>2021</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>1121</td>
<td>685</td>
<td>100.0% 0.37 [-0.10, 0.85]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Test for overall effect: Z = 1.53 (P = 0.13)

Heterogeneity: Tau² = 0.19; Chi² = 19.01, df = 3 (P = 0.0003); I² = 84%

1.1 within individuals

| Licht 2008         | 2.82 1.45 170      | 3.09 1.56 170   | 46.9% 0.27 [-0.59, 0.05] | 2008                  |
| Gao 2020           | 3.764 1.551 102    | 3.577 1.438     | 102 35.0% 0.19 [-0.32, 0.60] | 2020                  |
| Sacha 2021         | 2.9 1.5 42         | 3 1.5 42        | 18.1% 0.10 [-0.74, 0.54] | 2021                  |
| Subtotal (95% CI)  | 314                | 314              | 100.0% 0.08 [-0.38, 0.22] |                       |

Test for overall effect: Z = 0.52 (P = 0.61)

Test for subgroup differences: Chi² = 2.46, df = 1 (P = 0.12), I² = 59.3%

Figure 4B. comparison in sperm concentration

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Home collected Mean</th>
<th>Home collected SD</th>
<th>Home collected Total</th>
<th>Clinic collected Mean</th>
<th>Clinic collected SD</th>
<th>Clinic collected Total</th>
<th>Mean Difference</th>
<th>Weight</th>
<th>IV, Random, 95% CI</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall samples</td>
<td>58 40 236</td>
<td>59 40 397</td>
<td>30.1% -1.00 [-7.44, 5.44]</td>
<td>2007</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Song 2007</td>
<td>125.5 73.835 106</td>
<td>113.55 65.56</td>
<td>273 12.3% 11.95 [-4.11, 28.01]</td>
<td>2008</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gao 2020</td>
<td>60.7 33 837</td>
<td>51.9 36.9</td>
<td>244 33.5% 8.80 [3.66, 13.94]</td>
<td>2021</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sacha 2021</td>
<td>79.3 64.4 125</td>
<td>66.1 43.8</td>
<td>119 15.1% 13.20 [-0.56, 26.96]</td>
<td>2021</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>1357</td>
<td>1082</td>
<td>100.0% 5.46 [-1.23, 12.15]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Heterogeneity: Tau² = 27.92; Chi² = 8.90, df = 4 (P = 0.06); I² = 55%

Test for overall effect: Z = 2.16 (P = 0.11)

1.2 within individuals

| Licht 2008         | 51.29 44.49 170    | 52.85 48.81      | 170 63.5% -1.56 [-11.49, 8.37] | 2008                  |
| Sacha 2021         | 81 51.5 42         | 78.6 40.9        | 42 15.8% 2.46 [-17.49, 22.28] | 2021                  |
| Subtotal (95% CI)  | 314                | 314              | 100.0% 0.82 [-7.09, 8.73] |                       |

Heterogeneity: Tau² = 0.00; Chi² = 0.71, df = 2 (P = 0.70); I² = 0%

Test for overall effect: Z = 2.20 (P = 0.04)

Test for subgroup differences: Chi² = 0.77, df = 1 (P = 0.38), I² = 0%
Figure 4C. comparison in total sperm count

![Comparison in total sperm count](image)

Figure 4D. comparison in sperm motility

![Comparison in sperm motility](image)

Figure 4E. comparison in total motile sperm count

![Comparison in total motile sperm count](image)
Figure 4F. comparison in progressive sperm motility

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Home collected</th>
<th>Clinic collected</th>
<th>Mean Difference IV, Random, 95% CI</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Song 2007</td>
<td>30</td>
<td>7</td>
<td>236</td>
<td></td>
</tr>
<tr>
<td>Elzanaty 2008</td>
<td>49</td>
<td>18.26</td>
<td>106</td>
<td></td>
</tr>
<tr>
<td>Gao 2020</td>
<td>41.816</td>
<td>15.388</td>
<td>53</td>
<td></td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>395</td>
<td></td>
<td></td>
<td>2007</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Heterogeneity: Tau² = 6.74; Chi² = 6.48, df = 2 (P = 0.04); I² = 69%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Test for overall effect: Z = 0.64 (P = 0.52)</td>
<td></td>
</tr>
<tr>
<td>Elzanaty 2008</td>
<td>42.6</td>
<td>14.479</td>
<td>102</td>
<td></td>
</tr>
<tr>
<td>Gao 2020</td>
<td>41.579</td>
<td>14.414</td>
<td>102</td>
<td></td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>102</td>
<td></td>
<td></td>
<td>2020</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Heterogeneity: Not applicable</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Test for overall effect: Z = 0.50 (P = 0.61)</td>
<td></td>
</tr>
</tbody>
</table>

Test for subgroup differences: Chi² = 0.00, df = 1 (P = 0.99), I² = 0%

---

Figure 4G. comparison in normal morphology

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Home collected</th>
<th>Clinic collected</th>
<th>Mean Difference IV, Random, 95% CI</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elzanaty 2008</td>
<td>6.7</td>
<td>2.779</td>
<td>106</td>
<td></td>
</tr>
<tr>
<td>Gao 2020</td>
<td>3.395</td>
<td>2.131</td>
<td>53</td>
<td></td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>8.7</td>
<td></td>
<td></td>
<td>2006</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Heterogeneity: Tau² = 0.06; Chi² = 2.66, df = 2 (P = 0.26); I² = 25%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Test for overall effect: Z = 0.34 (P = 0.73)</td>
<td></td>
</tr>
<tr>
<td>Licht 2008</td>
<td>4.69</td>
<td>3.36</td>
<td>170</td>
<td></td>
</tr>
<tr>
<td>Gao 2020</td>
<td>3.905</td>
<td>1.974</td>
<td>102</td>
<td></td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>8.7</td>
<td></td>
<td></td>
<td>2020</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Heterogeneity: Tau² = 0.00; Chi² = 0.02, df = 1 (P = 0.65); I² = 0%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Test for overall effect: Z = 0.86 (P = 0.40)</td>
<td></td>
</tr>
</tbody>
</table>

Test for subgroup differences: Chi² = 0.45, df = 1 (P = 0.50), I² = 0%
Figure 5. Fertility outcomes

Figure 5A. Comparison in fertilization rate

Figure 5B. Comparison in pregnancy rate (IUI/IVF outcomes)
Figure 5. Fertility outcomes between home and clinic semen collection

Figure 5A. comparison in fertilization rate

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Home collected</th>
<th>Clinic collected</th>
<th>Risk Ratio M-H, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sacha 2021</td>
<td>859</td>
<td>1141</td>
<td>1.00 [0.96, 1.05]</td>
</tr>
<tr>
<td>Stimpfel 2021</td>
<td>3803</td>
<td>5968</td>
<td>1.00 [0.96, 1.04]</td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td>4692</td>
<td>2078</td>
<td>1.00 [0.97, 1.03]</td>
</tr>
</tbody>
</table>

Heterogeneity: $\tau^2 = 0.00; \chi^2 = 0.00, df = 1 (P = 0.95); I^2 = 0\%$

Test for overall effect: $Z = 0.20 (P = 0.84)$

Figure 5B. comparison in pregnancy rate (IUI/IVF outcomes)

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Home collected</th>
<th>Clinic collected</th>
<th>Risk Ratio M-H, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.2.1 IUI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yavas 2004</td>
<td>12</td>
<td>95</td>
<td>0.47 [0.22, 0.99] 2004</td>
</tr>
<tr>
<td>Song 2007</td>
<td>25</td>
<td>236</td>
<td>1.45 [0.87, 2.42] 2007</td>
</tr>
<tr>
<td>Total events</td>
<td>37</td>
<td>39</td>
<td>0.85 [0.28, 2.59]</td>
</tr>
</tbody>
</table>

Heterogeneity: $\tau^2 = 0.54; \chi^2 = 6.05, df = 1 (P = 0.01); I^2 = 83\%$

Test for overall effect: $Z = 0.28 (P = 0.78)$

| 2.2.2 IVF/ICSI    |                |                  |                                |
| Sacha 2021        | 38             | 67               | 1.17 [0.84, 1.63] 2021         |
| Stimpfel 2021     | 221            | 653              | 0.98 [0.79, 1.23] 2021         |
| Subtotal (95% CI) | 259            | 97               | 1.04 [0.86, 1.25]             |

Heterogeneity: $\tau^2 = 0.00; \chi^2 = 0.75, df = 1 (P = 0.39); I^2 = 0\%$

Test for overall effect: $Z = 0.42 (P = 0.66)$

Test for subgroup differences: $\chi^2 = 0.12, df = 1 (P = 0.73), I^2 = 0\%$