Temporal trends in sperm count: a systematic review and meta-regression analysis of samples collected globally in the 20th and 21st centuries

Hagai Levine 1,*, Niels Jørgensen 2,3, Anderson Martino-Andrade 4, Jaime Mendiola 5, Dan Weksler-Derri 6,7, Maya Jolles 1, Rachel Pinotti 8, and Shanna H. Swan 9

1Braun School of Public Health and Community Medicine, Hadassah Medical Center, The Faculty of Medicine, Hebrew University of Jerusalem, Jerusalem, Israel 2Department of Growth and Reproduction, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark 3International Center for Research and Research Training in Endocrine Disruption of Male Reproduction and Child Health (EDMaRC), Rigshospitalet, University of Copenhagen, Copenhagen, Denmark 4Department of Physiology, Federal University of Paraíba, Curitiba, Brazil 5Division of Preventive Medicine and Public Health, University of Murcia School of Medicine and Biomedical Research Institute of Murcia (IMIB-Arrixaca-UMU), Murcia, Spain 6Clalit Health Services, Kiryat Ono, Israel 7Faculty of Health Sciences, Ben-Gurion University of the Negev, Beer Sheva, Israel 8Gustave L. and Janet W. Levy Library, Icahn School of Medicine at Mount Sinai, New York, NY, USA 9Department of Environmental Medicine and Public Health, Icahn School of Medicine at Mount Sinai, New York, NY, USA

*Correspondence address. Braun School of Public Health and Community Medicine, Hadassah Medical Center, The Faculty of Medicine, Hebrew University of Jerusalem, Ein Kerem Campus, POB 12272, Jerusalem 91102, Israel. Tel: +972-50-517-2895; E-mail: hagai.levine@gmail.com

Submitted on June 10, 2022; resubmitted on September 29, 2022; editorial decision on October 11, 2022

TABLE OF CONTENTS

• Introduction
• Methods
  Systematic review and study selection
  Data extraction
  Quality control
  Statistical analysis
• Results
  Systematic review and summary statistics
  Simple linear models
  Meta-regression models for SC
  Meta-regression models for TSC
  Meta-regression models for recent periods
  Sensitivity analyses
• Discussion
  Key findings
  Comparison to prior studies
  Other issues
  Strengths
  Limitations
• Conclusion and wider implications
BACKGROUND: Numerous studies have reported declines in semen quality and other markers of male reproductive health. Our previous meta-analysis reported a significant decrease in sperm concentration (SC) and total sperm count (TSC) among men from North America–Europe–Australia (NEA) based on studies published during 1981–2013. At that time, there were too few studies with data from South/Central America–Asia–Africa (SAA) to reliably estimate trends among men from these continents.

OBJECTIVE AND RATIONALE: The aim of this study was to examine trends in sperm count among men from all continents. The broader implications of a global decline in sperm count, the knowledge gaps left unfilled by our prior analysis and the controversies surrounding this issue warranted an up-to-date meta-analysis.

SEARCH METHODS: We searched PubMed/MEDLINE and EMBASE to identify studies of human SC and TSC published during 2014–2019. After review of 2936 abstracts and 868 full articles, 44 estimates of SC and TSC from 38 studies met the protocol criteria. Data were extracted on semen parameters (SC, TSC, semen volume), collection year and covariates. Combining these new data with data from our previous meta-analysis, the current meta-analysis includes results from 223 studies, yielding 288 estimates based on semen samples collected 1973–2018. Slopes of SC and TSC were estimated as functions of sample collection year using simple linear regression as well as weighted meta-regression. The latter models were adjusted for predetermined covariates and examined for modification by fertility status (unselected by fertility versus fertile), and by two groups of continents: NEA and SAA. These analyses were repeated for data collected post-2000. Multiple sensitivity analyses were conducted to examine assumptions, including linearity.

OUTCOMES: Overall, SC declined appreciably between 1973 and 2018 (slope in the simple linear model: –0.87 million/ml/year, 95% CI: –0.89 to –0.86; P < 0.001). In an adjusted meta-regression model, which included two interaction terms [time × fertility group (P = 0.012) and time × continents (P = 0.058)], declines were seen among unselected men from NEA (–1.27; –1.78 to –0.77; P < 0.001) and unselected men from SAA (–0.65; –1.29 to –0.01; P = 0.045) and fertile men from NEA (–0.50; –1.00 to –0.01; P = 0.046). Among unselected men from all continents, the mean SC declined by 51.6% between 1973 and 2018 (–1.16% per year, –1.66 to –0.68; P < 0.001). The per- cent decline per year doubled, increasing from 1.16% post-1972 to 2.64% post-2000. Results were similar for TSC, with a 62.3% overall decline among unselected men (–4.70 million/year; –6.56 to –2.83; P < 0.001) in the adjusted meta-regression model. All results changed only minimally in multiple sensitivity analyses.

WIDER IMPLICATIONS: This analysis is the first to report a decline in sperm count among unselected men from South/Central America–Asia–Africa, in contrast to our previous meta-analysis that was underpowered to examine those continents. Furthermore, data suggest that this world-wide decline is continuing in the 21st century at an accelerated pace. Research on the causes of this continuing decline and actions to prevent further disruption of male reproductive health are urgently needed.
**Key words:** human reproduction / male infertility / andrology / semen quality / sperm count / semen analysis / environmental effects / epidemiology / systematic review / meta-analysis

### Introduction

In 2017, ‘Temporal Trends in Sperm Count: A Systematic Review and Meta-Regression Analysis’ was published by this journal (Levine et al., 2017). That article, which was widely discussed and highly cited, includes all eligible English-language publications in 1981–2013 that contained data on sperm count. We reported a very strong decline in sperm concentration (SC) and total sperm count (TSC) in North America, Europe, Australia/New Zealand (hereafter NEA) but too few studies have been published in South/Central America, Asia and Africa (hereafter SAA) to draw a conclusion about trends in those continents. We examined mean SC and TSC as a function of collection year, as approximated by the mid-year of the sample collection period. Because sample collection preceded the year of publication by an average of 6 years, our results were already somewhat dated by the time we published our analysis in 2017. Therefore, we conducted a new literature search in the spring of 2020 to identify eligible studies published between 1 January 2014 and 31 December 2019. Here, we report on global trends in SC and TSC in publications 1981 through 2019, which combines results of both searches and analyses. This expanded analysis addresses two important questions. With increased sample size, was a trend seen in South America, Africa and Asia? Did the trends we reported continue post-2011?

Recognition of the importance of male reproductive function, and sperm count, has increased since 2017 (Levine et al., 2018; United Nations Human Rights Office of the High Commissioner, 2019). The economic and societal burden of male infertility is now widely recognized, as is the unequal burden of male infertility which falls most heavily on low-income populations (Winters and Walsh, 2014; Hauser et al., 2015; Dupree, 2016; Skakkebæk et al., 2016). Increasingly strong evidence links reduced sperm count and concentration to increases in all-cause mortality and morbidity (Latif et al., 2017; Del Giudice et al., 2021; Ferlin et al., 2021). Links between sperm count and infertility are well-recognized (Bonde et al., 1998; Skakkebæk et al., 2022). Furthermore, the decline in sperm count is paralleled by declines in testosterone and increases in testicular cancer and male genital anomalies (Skakkebæk et al., 2022). In fact, the decline in semen quality and male reproductive health has recently been described as a crisis (De Jonge and Barratt, 2019). Relative to declines in sperm counts, these latter trends are far more difficult to document. There is currently no systematic collection of such data, making the examination of those trends difficult. Therefore, an international group of scientists, including several of the authors, has suggested the formation of a multidisciplinary monitoring system for reproductive health indicators that would provide ongoing surveillance of reproductive health outcomes (Le Moal et al., 2016).

The broad implications of a global decline in sperm count, the data gaps left unfilled by our prior analysis and the controversies surrounding these issues warrant an up-to-date meta-analysis. This meta-analysis was conducted to address these aims.

### Methods

This systematic review and meta-regression analysis was conducted, and the results reported, in accordance with Meta-analysis in Observational Studies in Epidemiology (Stroup et al., 2000) (Supplementary Table SI) and Preferred Reporting Items for Systematic reviews and Meta-Analysis guidelines (Liberati et al., 2009; Moher et al., 2009). Our research team included epidemiologists, andrologists and a medical librarian. Our predefined protocol was developed following best practices (Borenstein et al., 2009; Higgins and Green 2011; National Toxicology Program 2015) and informed by our previous study. Throughout, unless otherwise noted, the methods of the current study are those employed and published in the previous study (Levine et al., 2017), including keywords and databases searched, eligibility criteria and statistical methods.

**Systematic review and study selection**

A comprehensive search of the PubMed/MEDLINE and EMBASE was conducted, to identify English-language publications that reported primary data on human sperm count, published in 2014–2019, i.e. from the last date included in our prior search through 2019. On 15 May 2020, we searched MEDLINE/PubMed and Embase (Excerpta Medica database) for peer-reviewed publications meeting our inclusion criteria. Following the recommendation of the Cochrane Handbook for Systematic Reviews, we searched in title and abstract for both index (MeSH) terms and keywords and filtered out animal-only studies. We used the MeSH term ‘sperm count’, which includes seven additional terms, and to increase sensitivity, we added 13 related keywords (e.g. ‘sperm density’, ‘sperm concentration’).

All English-language studies that reported primary data on human SC were considered eligible for abstract screening. We evaluated the eligibility of all subgroups within a study. For example, in a case–control study, the control group might have been eligible for inclusion even though, based on our exclusion criteria, the case group was not.

We divided eligible studies into two fertility-defined groups: (i) studies of men unselected by fertility status, hereafter ‘Unselected’ (e.g. young men unlikely to be aware of their fertility, such as young men being screened for military service or college students) and (ii) studies of men whose partners had born a child or whose partners were pregnant regardless of pregnancy outcome, hereafter ‘Fertile’. ‘Total’ refers to both groups.

A study was excluded if participants were selected based on: (i) infertility or sub-fertility; (ii) their range of semen parameters (e.g. studies selecting normospermic men); and (iii) genital abnormalities, diseases or medications. We also excluded studies limited to men with exposures that may affect fertility, such as an occupational exposure, post-clinical trial intervention or smoking. Studies of candidates for vasectomy or semen donation were included only if semen quality was not a criterion for men’s study participation. We excluded studies that used non-standard methods for sperm collection (e.g. methods
other than masturbation) or counting methods other than hemocytometer and studies with fewer than 10 men.

First, the publication was either excluded or advanced to full-text screening based on the title and abstract. Any publication without an abstract was automatically referred for full-text screening. Second, we reviewed the full text and assigned it to exclusion (and categorized the reason for exclusion) or to data extraction. We then confirmed study eligibility and identified multiple publications from the same study to ensure that estimates from the same population were not used more than once.

**Data extraction**

For each estimate, we extracted summary statistics on SC and TSC (mean, standard deviation, minimum, maximum, median and percentiles), semen volume (mean and measured method), sample size, sample collection years, data on covariates (fertility group, country, age, ejaculation abstinence time reported and per protocol, method of semen collection, method of assessing of SC and semen volume, selection of population, study exclusion criteria and number of samples per man). The range of permissible values, both for categorical and numerical variables and information on data completeness were recorded. Data were extracted on the total population as well as on all eligible subgroups.

**Quality control**

The current study followed the same protocol as used in the previous study. For the new search, we used Covidence systematic review software for the process of screening the articles that were not available for our prior search. In addition, one member of the research team was replaced. Screening of this extensive systematic review was conducted by a team of eight reviewers (H.L., N.J., A.M.-A., J.M., D.W.-D., M.J., R.P., S.H.S.). As in our previous analysis, the screening protocol was piloted and reviewers were trained by screening of 50 abstracts by all reviewers followed by a comparison of results, resolution of any inconsistencies and clarification of the protocol as needed. The same quality control process was followed for full-text screening and data extraction by all reviewers. All data were entered into digital spreadsheets with explicit permissible values (no open-ended entries) to increase consistency. After data extraction, an additional round of data editing and quality control of all studies was conducted by H.L. and M.J. The process ensured that each study was evaluated by at least two different trained reviewers.

**Statistical analysis**

We ran all models both on the data used in our 2017 study (for quality control) and for the dataset including all years. In all models, the midpoint of the sample collection period was the independent variable (‘collection year’) and mean SC (or TSC) was the dependent variable(s). Units were million/ml for SC and million for TSC (reported or defined as SC × sample volume). All slopes denote unit change per calendar year, reported with 95% confidence intervals (CI). We also reported values for the first and last years, and the % change/year.

We first used simple linear regression models to estimate SC and TSC as functions of collection year, with each study weighted by sample size. Beside a model for all men (Unselected plus Fertile), we also ran a model stratified by fertility status.

Then, we used random-effects meta-regression to model both SC and TSC as linear functions of time, weighting by the standard error (SE).

In the meta-regression models, we controlled for a predetermined set of potential confounders: fertility group, age, abstinence time, whether semen collection and counting methods were reported, number of samples per man and indicators for exclusion criteria (Supplementary Table SII). The method used to determine sample volume was extended when modeling TSC. In all meta-regression analyses, we included indicator variables to denote studies with more than one SC/TSC estimate.

Several variables were imputed including mean SC (and TSC), as described in our 2017 study. For example, for studies that reported median (not mean) SC or TSC, we estimated the mean by adding the range difference between the mean and median in studies for which both were reported.

We included indicators in all models for imputed values. We included continental group (NEA or SAA) and fertility group (Unselected or Fertile), as variables in the model when applicable.

We ran several meta-regression models for both SC and TSC:

1. basic, unadjusted model;
2. adjusted model for all men; and
3. stratified by fertility; and
4. unselected men only with time × continent interaction; and
5. all men with two interactions: time × fertility and time × continent.

These analyses were repeated for subsets of data collection to examine recent trends (>1972, >1985, >1995 and >2000).

We conducted several sensitivity analyses: (i) adding cubic and quadratic terms for collection year in meta-regression analyses to assess non-linearity; (ii) removing covariates one at a time from the model; (iii) excluding a specific group, i.e. the group with no information on age, for each covariate; (iv) replacing age group by mean age, excluding studies that did not report mean age; (v) adding a covariate for high smoking prevalence (>30%); (vi) removing each continent one at a time; (vii) excluding studies with five or more data points to examine the influence of large studies; and (viii) removing studies with SEs >20 million/ml.

All analyses were conducted using STATA version 14.1 (StataCorp).

**Results**

**Systematic review and summary statistics**

Using PubMed/MEDLINE and Embase searches, we identified 2936 new publications meeting our criteria for abstract screening (Fig. 1). Of these, 151 duplicate records were removed and 1917 were excluded based on title or abstract screening. Full texts of the remaining 868 articles were reviewed for eligibility and 743 studies were excluded. Of the remaining 125 articles, 87 were excluded during data extraction through the second round of full-text screening; 44 of them were duplicate publications. The remaining 38 studies of semen samples from 14 233 men included 44 unique mean SC estimates that met the protocol criteria. Combining these new data with data from our previous meta-analysis, the current meta-analysis includes results from 223 studies, yielding 288 estimates based on semen samples collected...
Records identified through database searching (n=2936)
PubMed/MEDLINE:2161
Embase:775

Title and abstract screening (n=2936)
- Duplicates removed (n=151)
- Title and abstract screening (n=1917)

1st round of Full-text (FT) screening (n=868)

Exclusion based on 2nd round of FT screening (n=87)
- Selected by Infertility: 9
- Excluded genital abnormality or infertility: 4
- Sperm count method other than hemocytometer: 9
- No sperm count data: 8
- Selected by Semen: 6
- Selected by Disease: 2
- Sperm count method missing: 2
- No original data: 1
- Selection missing: 2

288 estimates (153 unselected by fertility, 135 fertile)

44 estimates (13 unselected by fertility, 31 fertile)

Exclusion based on 1st round of FT screening (n=743)
- Selected by Infertility: 148
- Excluded genital abnormality or infertility: 115
- Sperm count method other than hemocytometer: 101
- No sperm count data: 100
- Selected by Semen: 100
- Selected by Disease: 34
- Sperm count method missing: 15
- No original data: 45
- Selection missing: 9
- Less than 10 men: 13
- Not English: 5
- Selected by Intervention: 5
- Not humans: 4
- Selected by Exposure: 3
- Not masturbation: 2
- Selected by Smoking: 1
- No full text: 43

Previous: 244 estimates (140 unselected by fertility, 104 fertile)

Figure 1. PRISMA flow chart showing the selection of studies eligible for meta-regression analysis.

1973–2018 provided by 57,168 men. Data were available from 6 continents and 53 countries (complete list in Supplementary Table SIII and Supplementary Fig. S1). Of the 288 estimates, 118 (41%) were Unselected NEA, 35 (12%) were Unselected SAA, 81 (28%) were Fertile NEA and 54 (19%) were Fertile SAA. The number of estimates from SAA increased by 29%.
Simple linear models

Combining results from all men, SC declined steeply (slope per year -0.87 million/ml; 95% CI: -0.89 to -0.86; P < 0.001) between 1973 and 2018 when using simple linear models, unadjusted and weighted by sample size (Supplementary Table SIV, Supplementary Fig. S2). For all men combined, SC declined by 0.93% per year and overall, by 41.5% between 1973 and 2018. In a model stratified by fertility group, the slope was much steeper for Unselected (–1.23; –1.25 to –1.20; P < 0.001) than for Fertile men (–0.30; –0.33 to –0.27; P < 0.001) (Supplementary Table SIV). A similar trend was seen for TSC when combining the two fertility groups (slope per year = –2.80 million; –2.86 to –2.74; P < 0.001), and the slope was steeper for the Unselected group (–3.77; –3.83 to –3.71; P < 0.001) (Supplementary Table SIV, Supplementary Fig. S2). Semen volume did not change over the study period (slope per year = 0.0002 ml; –0.0001 to 0.0005; P = 0.249).

Meta-regression models for SC

In a basic meta-regression model for SC, in which estimates were weighted by their SE but without covariate adjustment, slopes were slightly less steep than for the simple regression model, and with wider CIs (slope per year = –0.66 million/ml; 95% CI: –0.92 to –0.40; P < 0.001). Covariate adjustment did not appreciably alter the slope but widened the CI further (–0.59; –0.90 to –0.27; P < 0.001) (Table I, betas for covariates in Supplementary Table SII).

After stratifying by fertility group and adjusting for all covariates including continental group, there was a strong decline in SC among unselected men (–1.17; –1.66 to –0.68; P < 0.001) but not among fertile men (–0.11; –0.54 to 0.32; P = 0.615) (Table I, Fig. 2). Using SC model estimates of 101.2 million/ml in 1973 and 49.0 million/ml in 2020. (Table I, Fig. 3).

In an adjusted meta-regression model among unselected men that included interaction by geographic group (P for interaction = 0.44), the slope for Unselected NEA was –1.30 (–1.89 to –0.71; P < 0.001) and the slope for Unselected SAA was –0.84 (–1.82 to 0.13; P = 0.088) (Table I, Fig. 3).

In an adjusted meta-regression model, which included all men and two interaction terms [time × fertility group (P = 0.012) and time × continents (P = 0.058)], declines were seen among Unselected NEA...
men (–4.70; –6.56 to –2.83). In an adjusted meta-regression model, which included all men and two interaction terms [time × fertility group (P = 0.013) and time × continents (P = 0.015)], declines were seen among Unselected NEA (–4.71; –6.53 to –2.88; P < 0.001), Unselected SAA (–1.78; –4.10 to 0.55; P = 0.133) and Fertile NEA (–2.09; –3.86 to –0.32; P = 0.021) (Table I, Supplementary Fig. S3).

**Meta-regression models for recent periods**

We also restricted the analysis of unselected men (in all continents) to recent time intervals (Table II, Fig. 2). Post-1995, the slope for SC was somewhat steeper (–1.33; –2.41 to –0.26; P = 0.016) and was steeper still post-2000 (–1.73; –3.23 to –0.24; P = 0.024). There was a marked increase in the percent decline in SC per year in the recent period, from 1.16% post-1972 to 2.64% post-2000 (Fig. 4). Post-2000, the slope for TSC (–5.26, –10.72 to 0.19; P = 0.058) was also steeper than that for post-1972 (Table II, Fig. 2).

**Sensitivity analyses**

We performed multiple analyses to examine the sensitivity of results to assumptions about our model, linearity, influence of covariates and

### Table I Sperm concentration (SC) and total sperm count (TSC) in first and last years of meta-regression analysis, adjusted for continents and potential confounders,* with % change and slope per year: (i) total; (ii) stratified by fertility; (iii) unselected men only with time × continent interaction; (iv) two interactions: time × fertility and time × continent.

<table>
<thead>
<tr>
<th>Model</th>
<th>Category</th>
<th>N (estimates)</th>
<th>First Year</th>
<th>Last Year</th>
<th>%change/year</th>
<th>Slope (95% CI), million/ml/year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>All men</td>
<td>288</td>
<td>1973</td>
<td>2018</td>
<td>–0.71</td>
<td>–0.59 (–0.90 to –0.27)</td>
</tr>
<tr>
<td></td>
<td>Unselected</td>
<td>153</td>
<td>1973</td>
<td>2018</td>
<td>–1.16</td>
<td>–1.17 (–1.66 to –0.68)</td>
</tr>
<tr>
<td></td>
<td>Fertile</td>
<td>135</td>
<td>1977</td>
<td>2017</td>
<td>–0.14</td>
<td>–0.11 (–0.54 to 0.32)</td>
</tr>
<tr>
<td>Stratified</td>
<td>Unselected NEAb</td>
<td>118</td>
<td>1973</td>
<td>2015</td>
<td>–1.25</td>
<td>–1.30 (–1.89 to –0.71)</td>
</tr>
<tr>
<td></td>
<td>Unselected SAAb</td>
<td>35</td>
<td>1986</td>
<td>2018</td>
<td>–0.96</td>
<td>–0.84 (–1.82 to 0.13)</td>
</tr>
<tr>
<td></td>
<td>Fertile NEAb</td>
<td>81</td>
<td>1977</td>
<td>2017</td>
<td>–0.59</td>
<td>–0.50 (–1.00 to –0.01)</td>
</tr>
<tr>
<td></td>
<td>Fertile SAAb</td>
<td>54</td>
<td>1978</td>
<td>2016</td>
<td>0.18</td>
<td>0.13 (–0.42 to 0.67)</td>
</tr>
</tbody>
</table>

*Meta-regression model, adjusted for continents, age, abstinence time, semen collection method reported, counting method reported, having more than one sample per men, indicators for study selection of population and exclusion criteria (some vasectomy candidates, some semen donor candidates, exclusion of men with chronic diseases, exclusion by other reasons not related to fertility, selection by occupation not related to fertility), whether collection year was estimated, whether arithmetic mean of SC was estimated, whether SE of SC was estimated and indicator variable to denote studies with more than one estimate. Sperm concentration (SC) meta-regression models weighted by SC SE, adjusted for similar covariates and method used to assess semen volume. Total sperm count (TSC) meta-regression models weighted by TSC SE, adjusted for similar covariates and method used to assess semen volume.

bNEA, North America–Europe–Australia; SAA, South/Central America–Asia–Africa.

(–1.27; –1.78 to –0.77; P < 0.001), Unselected SAA (–0.65; –1.29 to –0.01; P = 0.045) and Fertile NEA (–0.50; –1.00 to –0.01; P = 0.046) (Table I, Supplementary Fig. S3).

**Meta-regression models for TSC**

Overall, TSC trends were similar to those for SC. In an adjusted meta-regression model for all men, there was a steep decline in TSC (slope per year –2.06 million, –3.25 to –0.87; P = 0.001) (Table I).

After stratifying by fertility group and adjusting for all covariates including continent, there was a strong decline in TSC among unselected men (–4.70; –6.56 to –2.83; P < 0.001) but not among fertile men (–0.24; –1.99 to 1.52; P = 0.788) (Table I, Fig. 2). Using TSC model estimates of 335.7 million in 1973 and 126.6 million in 2018, TSC declined among unselected men by 1.40% per year and 62.3% overall (Table I).

In an adjusted meta-regression model among unselected men, including interaction by geographic group (P for interaction = 0.44), the slope for Unselected NEA was –5.05 (–7.31 to –2.79; P < 0.001) and the slope for Unselected SAA was –3.79 (–7.58 to –0.01, P = 0.049) (Table I, Fig. 3).
Figure 2. Meta-regression models for mean sperm concentration (SC) and total sperm count (TSC) by collection year among unselected men from all continents, adjusted for potential confounders, for the whole period and restricted to studies post 2000. (a) Sperm concentration. (b) Total sperm count. Meta-regression model weighted by sperm concentration (SC) SE, adjusted for continents, age, abstinence time, semen collection method reported, counting method reported, having more than one sample per man, indicators for study selection of population and exclusion criteria (some vasectomy candidates, some semen donor candidates, exclusion of men with chronic diseases, exclusion by other reasons not related to fertility, selection by occupation not related to fertility), whether collection year was estimated, whether arithmetic mean of SC was estimated, whether SE of SC was estimated and indicator variable to denote studies with more than one estimate. Total sperm count (TSC) meta-regression models weighted by TSC SE, adjusted for similar covariates and method used to assess semen volume. SE, standard error.

Figure 3. Meta-regression models for mean sperm concentration (SC) and total sperm count (TSC) by collection year with interaction for continents among unselected men, adjusted for potential confounders. (a) Sperm concentration (SC). (b) Total sperm count. NEA, North America–Europe–Australia; SAA, South/Central America–Asia–Africa. Meta-regression model weighted by sperm concentration (SC) SE, adjusted for continents, age, abstinence time, semen collection method reported, counting method reported, having more than one sample per man, indicators for study selection of population and exclusion criteria (some vasectomy candidates, some semen donor candidates, exclusion of men with chronic diseases, exclusion by other reasons not related to fertility, selection by occupation not related to fertility), whether collection year was estimated, whether arithmetic mean of SC was estimated, whether SE of SC was estimated and indicator variable to denote studies with more than one estimate. Total sperm count (TSC) meta-regression models weighted by TSC SE, adjusted for similar covariates and method used to assess semen volume. SE, standard error.
Table II  Stratified meta-regression model\textsuperscript{a} for mean sperm concentration (SC) and mean total sperm count (TSC) among unselected men, by periods.

<table>
<thead>
<tr>
<th>Period (years)</th>
<th>N (estimates)</th>
<th>First year</th>
<th>First year SC (million/ml)</th>
<th>Last year</th>
<th>Last year SC (million/ml)</th>
<th>% change/year</th>
<th>Slope (95% CI), million/ml/year</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;1972</td>
<td>153</td>
<td>1973</td>
<td>101.2</td>
<td>2018</td>
<td>49.0</td>
<td>–1.16</td>
<td>–1.17 (–1.66 to –0.68)</td>
</tr>
<tr>
<td>&gt;1985</td>
<td>131</td>
<td>1985</td>
<td>82.3</td>
<td>2018</td>
<td>47.1</td>
<td>–1.31</td>
<td>–1.08 (–1.68 to –0.49)</td>
</tr>
<tr>
<td>&gt;1995</td>
<td>89</td>
<td>1995</td>
<td>70.1</td>
<td>2018</td>
<td>40.1</td>
<td>–1.90</td>
<td>–1.33 (–2.41 to –0.26)</td>
</tr>
<tr>
<td>&gt;2000</td>
<td>60</td>
<td>2000</td>
<td>65.6</td>
<td>2018</td>
<td>35.3</td>
<td>–2.64</td>
<td>–1.73 (–2.33 to –0.24)</td>
</tr>
</tbody>
</table>

\textsuperscript{a}Stratified meta-regression model, adjusted for continents, age, abstinence time, semen collection method reported, counting method reported, having more than one sample per man, indicators for study selection of population and exclusion criteria (some vasectomy candidates, some semen donor candidates, exclusion of men with chronic diseases, exclusion by other reasons not related to fertility, selection by occupation not related to fertility), whether collection year was estimated, whether arithmetic mean of SC was estimated, whether SE of SC was estimated and indicator variable to denote studies with more than one estimate. Sperm concentration (SC) meta-regression models weighted by SC SE, adjusted for similar covariates and method used to assess semen volume. Total sperm count (TSC) meta-regression models weighted by TSC SE, adjusted for similar covariates and method used to assess semen volume.

Figure 4. Percent of decline according to periods, for mean sperm concentration (SC) and total sperm count (TSC) among unselected men using stratified meta-regression model. Stratified meta-regression model weighted by sperm concentration (SC) SE, adjusted for continents, age, abstinence time, semen collection method reported, counting method reported, having more than one sample per man, indicators for study selection of population and exclusion criteria (some vasectomy candidates, some semen donor candidates, exclusion of men with chronic diseases, exclusion by other reasons not related to fertility, selection by occupation not related to fertility), whether collection year was estimated, whether arithmetic mean of SC was estimated, whether SE of SC was estimated and indicator variable to denote studies with more than one estimate. Sperm concentration (SC) stratified meta-regression models weighted by SC SE, adjusted for similar covariates and method used to assess semen volume. SE, standard error.
imputation of missing data. Results from sensitivity analyses are only shown here for SC in unselected men (slope = −1.17 million/ml/year in the main model). In all sensitivity analyses (except one elaborated hereafter), there was a strong (>1.0 million/ml/year) decline in SC in the Unselected group, with \( P < 0.01 \).

Adding a quadratic or cubic function of year to meta-regression model did not substantially change the association between year and SC or improve the model fit: coefficient for the quadratic term: 0.04; 95% CI: −0.07 to 0.08, \( P = 0.135 \); for the cubic term 0.0005; 95% CI: −0.0007 to 0.001, \( P = 0.086 \).

For each covariate, we conducted two sensitivity analyses: (i) removing the covariate and (ii) by excluding a specific group, for each covariate (Supplementary Table SV). Excluding 47 estimates with no data on mean age and adjusting for mean age instead of age group, yielded a slope of −1.27 million/ml/year (−1.86 to −0.68; \( P < 0.001 \)), which is similar to the main model. The sensitivity analysis which showed a more than minimal change in slope was the exclusion of 39 estimates with no information on age, which yielded a slightly diminished slope of −0.94 million/ml/year (−1.51 to −0.37, \( P = 0.002 \)).

The proportion of smokers was reported in only 26.0% of studies and in 18.1% of studies of unselected men. To examine this variable, we ran a sensitivity analysis including a covariate for ‘high proportion of smokers’ (> 30%), and the slopes changed only slightly (−1.20 million/ml, −1.70 to −0.71; \( P < 0.001 \)).

Results for Unselected men did not materially change with additional sensitivity analysis, by exclusion of estimates from any specific continent (Supplementary Table SVI). Slopes were also robust after excluding the four studies with five or more data points (−1.04, −1.55 to −0.53; \( P < 0.001 \)), or excluding five estimates with a SE of SC > 20 million/ml (−1.11, −1.61 to −0.62; \( P < 0.001 \)).

Due to a typo in the value extracted for Rubes et al. (2010) in the previous meta-analysis, we repeated the analysis without this study as well as with the corrected value. The results did not materially change.

Discussion
The results of the present study extend those of our 2017 meta-analysis. As further elaborated below, the new data allow for new analyses and new results. We provide strong evidence, for the first time, of a decline in sperm counts among men from South/Central America, Asia, and Africa, as well as a world-wide decline in the 21st century, with data suggesting that the pace of this decline has accelerated.

Key findings
In our prior systematic review and meta-analysis (Levine et al., 2017), we reported a marked, continuing decline in both SC and TSC in North America, Europe and Australia based on samples collected between 1973 and 2011. What is new in the current analysis?

Our current analysis, the largest ever to examine temporal trends in sperm counts, extends both the study period and the number of estimates. This new analysis includes seven additional years of sample collection and adds 44 estimates to the 244 included in the earlier analysis. It is therefore both more robust and more temporally relevant.

The distribution of contributing countries has changed since our 2017 analysis. The new analysis appreciably increases the number of studies from SAA. With this increase in sample size, there is now adequate power to examine trends in SC and TSC in those continents. This analysis provides strong evidence, for the first time, of an appreciable decline in sperm counts among unselected men from SAA. Importantly, this analysis also demonstrates an accelerated decline in SC and TSC post-2000. In summary, this update confirms, extends and strengthens the results of our 2017 analysis.

Comparison to prior studies
Table III compares basic characteristics and results of the current study with those of Carlsen et al. (1992), Swan et al. (2000) and Levine et al. (2017), studies that together include data collected over more than 80 years. It is notable that although search and statistical analysis methods have become more sophisticated, and the distribution of studies has changed (with the proportion from SAA increasing), these slopes are remarkably consistent.

Comparing the current analysis with Levine et al. (2017), we note that the methods for searching and screening the literature, which are well documented in both, have not changed, nor have the analytic methods.

In both our current and past analyses, we excluded studies that selected men based on criteria that were likely to affect sperm count (e.g. requiring a minimum sperm count, or men’s participation in a sperm bank) with one exception. Studies of fertile men were included as a separate stratum (denoted Fertile). This group of studies includes fathers or partners of pregnant women. Thus, these men had either themselves helped to conceive a pregnancy, or the pregnancy was the result of in vitro fertilization (IVF). The proportion of IVF births has increased over the study period of this analysis, with eight million IVF babies born worldwide since the world’s first IVF birth in 1978 (European Society of Human Reproduction and Embryology, 2018). Among the 135 studies categorized as Fertile in this analysis, only 27.4% explicitly excluded pregnancies conceived by IVF.

Here, as in our 2017 analysis, we stratified countries into two groups, because of the potential for confounding of trends by geography. In the past, we referred to these two groups of countries as ‘Western’ and ‘Other’. Though not our intent, it became apparent that these terms had the potential to be misinterpreted and become politicized. Therefore, we now refer to these two groups of studies by the continent in which data were collected: ‘NEA’ (North America, Europe, and Australia) and ‘SAA’ (South/Central America, Africa, and Asia). We present results stratified by continental group, as well as combined.

Other issues
Could the declines we report be simply the result of a random decrease in a pattern of fluctuations (termed ‘sperm variability’) (Boulicault et al., 2021)? The continued decline demonstrated in this, and earlier, meta-analyses provide strong evidence that this is not the case. If, in fact, the declines we are reporting were merely the result of random fluctuation in sperm count, we would expect, on average, some percent of studies to report no change and the remainder to
The methods we used for the systematic literature review and meta-analysis are the most current and widely accepted by the scientific community.

### Limitations

We analyzed sperm count (SC and TSC) but not sperm motility and morphology. Interpreting trends in sperm motility and morphology is difficult, since methods have changed markedly over the study period. However, methods for measuring SC have remained largely unchanged. Counting by hemocytometer is the classical way to assess SC and has been recommended by the World Health Organization in all versions of organizations semen analysis manuals (Wang et al., 2022).

If no counting method was stated or use of a method other than hemocytometer was reported, the study was excluded. Overall, 334 studies were excluded because a method other than hemocytometer was used, while 128 studies were excluded because no sperm count method was provided. Of the 288 estimates for SC and TSC included, about half (146) were from studies that stated explicitly that a hemocytometer was used. The remaining 142 estimates were from studies that used World Health Organization methods without naming the particular type of hemocytometer used. We included both groups of studies in our analyses together with a variable indicating whether hemocytometer had been named explicitly as the counting method. In the sensitivity analyses, none of the slopes changed appreciably if we restricted the analysis to studies in which hemocytometer was named as the counting method or if this indicator variable was removed from the model.

Complete elimination of all selection/recruitment bias is impossible, since it is not possible to collect semen samples at random. However, we minimized selection bias by evaluating recruitment methods in all included studies. As stated in Methods, studies which selected men based on a semen parameter, a condition associated with a semen parameter (such as varicocele), or an exposure or occupation associated with semen quality. However, we did include ‘Fertile’ men as a separate group, even though this is a selected group. Compared to unselected men, the slope for fertile men was more modest than that for unselected men from NEA, and no decline was seen among fertile men from SAA. Men classified as ‘fertile’ are problematic in several ways. First, they include

<table>
<thead>
<tr>
<th>Table III Characteristics and results of fitting a simple linear regression model (without adjustment, weighted by sample size) for trends of sperm concentration in the current study compared to previous studies.</th>
</tr>
</thead>
<tbody>
<tr>
<td>First author (publication year)</td>
</tr>
<tr>
<td>Number of studies</td>
</tr>
<tr>
<td>Number of countries</td>
</tr>
<tr>
<td>Fertility group: N (%)</td>
</tr>
<tr>
<td>Unselected</td>
</tr>
<tr>
<td>Fertile</td>
</tr>
<tr>
<td>Continents: N (%)</td>
</tr>
<tr>
<td>NEA</td>
</tr>
<tr>
<td>SAA</td>
</tr>
<tr>
<td>Slope</td>
</tr>
<tr>
<td>P-value</td>
</tr>
</tbody>
</table>

*Wife pregnant or post-partum or at least 90% of men with proven fertility.

*NEA includes studies from North America–Europe–Australia. SAA includes studies from South/Central America–Asia–Africa.

While sperm count is an imperfect proxy for fertility, SC and TSC are closely linked to fertility chances (Guzick et al., 2001). The relationship between SC and time to conception is nonlinear. Thus, past a threshold of 40–50 million/ml, a higher SC does not necessarily imply a higher probability of conception. On the other hand, below that threshold, the probability of conception drops off rapidly as SC declines (Bonde et al., 1998). On a population level, the drop in mean SC from 104 to 49 million/ml that we report here implies a substantial increase in the proportion of men with delayed time to conception. Thus, SC provides the most stable and reliable measurement for comparisons within and among populations and over time.

**Strengths**

In this study, as in Levine et al. (2017), we used written protocols and extensive quality control procedures that minimized potential information and selection bias in all steps of the study. Further strengths of this study include our complete and documented literature search, the review of all retrieved articles by two members of the study team, and the use of current meta-analytic methods. All estimates were weighted by the SE of the measurement and all assumptions were examined in sensitivity analyses. In this study, we re-ran all steps of the prior meta-analyses on the larger, combined data set, as well as on the newly retrieved publications.

Our large number of studies and data points allowed us to control for a pre-determined set of covariates as well as for modification by fertility status and geographic group and variables indicating data completeness and study exclusion criteria.

The methods we used for the systematic literature review and meta-analysis are the most current and widely accepted by the scientific community.

---

The literature does not support this (Jørgensen et al., 2021; Aitken, 2022).

The methods we used for the systematic literature review and meta-regression of trends in sperm count are the most current and widely accepted by the scientific community.

We analyzed sperm count (SC and TSC) but not sperm motility and morphology. Interpreting trends in sperm motility and morphology is difficult, since methods have changed markedly over the study period. However, methods for measuring SC have remained largely unchanged. Counting by hemocytometer is the classical way to assess SC and has been recommended by the World Health Organization in all versions of organizations semen analysis manuals (Wang et al., 2022).

If no counting method was stated or use of a method other than hemocytometer was reported, the study was excluded. Overall, 334 studies were excluded because a method other than hemocytometer was used, while 128 studies were excluded because no sperm count method was provided. Of the 288 estimates for SC and TSC included, about half (146) were from studies that stated explicitly that a hemocytometer was used. The remaining 142 estimates were from studies that used World Health Organization methods without naming the particular type of hemocytometer used. We included both groups of studies in our analyses together with a variable indicating whether hemocytometer had been named explicitly as the counting method. In the sensitivity analyses, none of the slopes changed appreciably if we restricted the analysis to studies in which hemocytometer was named as the counting method or if this indicator variable was removed from the model.

Complete elimination of all selection/recruitment bias is impossible, since it is not possible to collect semen samples at random. However, we minimized selection bias by evaluating recruitment methods in all included studies. As stated in Methods, studies which selected men based on a semen parameter, a condition associated with a semen parameter (such as varicocele), or an exposure or occupation associated with semen quality. However, we did include ‘Fertile’ men as a separate group, even though this is a selected group. Compared to unselected men, the slope for fertile men was more modest than that for unselected men from NEA, and no decline was seen among fertile men from SAA. Men classified as ‘fertile’ are problematic in several ways. First, they include...
both those whose partner has conceived without medical assistance and with medical assistance, the fraction of which varied by time and location. Second, men with lower semen quality are underrepresented among fertile men. Changes in the proportion of fertile men in the population over time could lead to a selection bias in the Fertile group. In contrast, the Unselected group is not prone to such selection bias.

It is also possible that men providing a semen sample differ from those who do not. We previously studied this important question by comparing testosterone and Inhibin B levels in unselected men (potential military recruits undergoing a compulsory routine physical exam to determine their fitness for military service) who agreed to deliver a semen sample to those in men who only agreed to give a blood sample. In both groups of men, the hormone levels were similar (Andersen et al., 2000). Therefore, recruitment bias is unlikely in studies of unselected men.

As in our prior analysis, we included only English-language publications, which was unavoidable given the size of the task and the limited size of our study team. However, of the 2936 publications identified through our database searches in 2020, only 49 were excluded because of language.

In addition, it would be interesting to explore trends in sperm count in a specific continent or even within countries and sub-populations. However, we had inadequate statistical power to examine this question at a finer geographic level. Repeated studies on semen quality in specific populations would complement the current study by providing information about local trends.

Conclusion and wider implications

Our new data and analyses confirm our prior findings of an appreciable decline in sperm count between 1973 and 2018 among men from North America, Europe and Australia and support a decline among unselected men from South/Central America, Africa and Asia. This decline has continued, as predicted by our prior analysis, and has become steeper since 2000. This substantial and persistent decline is now recognized as a significant public health concern. In 2018, a group of leading clinicians and scientists called for governments to acknowledge decreased male fertility as a major public health problem and to recognize the importance of male reproductive health for the survival of the human (and other) species (Levine et al., 2018). Research on the causes of this continuing decline and an immediate focused response to prevent further disruption of male reproductive health are needed.

We hope that the new evidence provided here will receive attention not only from clinicians and scientists, but also from decision-makers and the general public.

Supplementary data

Supplementary data are available at Human Reproduction Update online.

Data availability

The data underlying this article will be shared on reasonable request to the corresponding author.

Acknowledgements

We thank Alexandria Albert for her assistance in the literature review.

Authors’ roles


Funding

We gratefully acknowledge financial support from The Grantham Foundation for the Protection of the Environment. None of the funding sources played a role in the design and conduct of the study; collection, management, analysis and interpretation of the data; preparations, review or approval of the manuscript; or decision to submit the manuscript for publication. Anderson Martino-Andrade is a scholarship recipient from the National Council for Scientific and Technological Development (CNPq, Brazil).

Conflict of interest

All authors declare they have no conflict of interest.

References


de Gouveia Brazao CA, Pieratt CLR. The present crisis in male reproductive health: an urgent need for a political, social, and research road-map. *Andrology* 2019;7:762–768.


Huang LP, Lee CC, Fan JP, Kuo PH, Shih TS, Hsu PC. Urinary metabolites of di(2-ethylhexyl) phthalate relation to sperm motility, reactive oxygen species generation, and apoptosis in polyvinyl chloride workers. *Int Arch Occup Environ Health* 2014;87:635–646.


