The future of sperm: a biovariability framework for understanding global sperm count trends

Marion Boulicault, Meg Perret, Jonathan Galka, Alex Borsa, Annika Gompers, Meredith Reiches & Sarah Richardson

To cite this article: Marion Boulicault, Meg Perret, Jonathan Galka, Alex Borsa, Annika Gompers, Meredith Reiches & Sarah Richardson (2021): The future of sperm: a biovariability framework for understanding global sperm count trends, Human Fertility, DOI: 10.1080/14647273.2021.1917778

To link to this article: https://doi.org/10.1080/14647273.2021.1917778

Published online: 10 May 2021.
The future of sperm: a biovariability framework for understanding global sperm count trends

Marion Boulicault, Meg Perret, Jonathan Galka, Alex Borsa, Annika Gompers, Meredith Reiches and Sarah Richardson

Department of Philosophy and Linguistics, Massachusetts Institute of Technology, Cambridge, MA, USA; Department of Philosophy, University of Adelaide, Adelaide, Australia; Department of the History of Science, Harvard University, Cambridge, MA, USA; Department of Sociomedical Sciences, Columbia University, New York, NY, USA; Department of Obstetrics and Gynecology, Beth Israel Deaconess Medical Center, Boston, MA, USA; Department of Anthropology, University of Massachusetts, Boston, MA, USA; Committee on Degrees in Studies of Women, Gender, and Sexuality, Harvard University, Cambridge, MA, USA

ABSTRACT

The past 50 years have seen heated debate in the reproductive sciences about global trends in human sperm count. In 2017, Levine and colleagues published the largest and most methodologically rigorous meta-regression analysis to date and reported that average total sperm concentration among men from ‘Western’ countries has decreased by 59.3% since 1973, with no sign of halting. These results reverberated in the scientific community and in public discussions about men and masculinity in the modern world, in part because of scientists’ public-facing claims about the societal implications of the decline of male fertility. We find that existing research follows a set of implicit and explicit assumptions about how to measure and interpret sperm counts, which collectively form what we term the Sperm Count Decline hypothesis (SCD). Using the study by Levine and colleagues, we identify weaknesses and inconsistencies in the SCD, and propose an alternative framework to guide research on sperm count trends: the Sperm Count Biovariability hypothesis (SCB). SCB asserts that sperm count varies within a wide range, much of which can be considered non-pathological and species-typical. Knowledge about the relationship between individual and population sperm count and life-historical and ecological factors is critical to interpreting trends in average sperm counts and their relationships to health and fertility.

ARTICLE HISTORY
Received 1 October 2020
Accepted 5 March 2021

KEYWORDS
Sperm count decline; male reproductive health; male fertility; semen analysis; andrology; human reproduction

Introduction

‘Who’s Killing America’s Sperm?’ read a 2017 cover of Newsweek magazine (Walsh, 2017). ‘Within a generation, men may lose the ability to reproduce entirely,’ pronounced Gentleman’s Quarterly (Halpern, 2018). ‘Western nations – although not developing countries – appear to be facing disaster,’ added The Guardian (McKie, 2017).

The instigating event was the publication, in Human Reproduction Update, of a meta-analysis of global sperm count change over time (Levine et al., 2017). The authors reported that total sperm count and sperm concentration (hereafter ‘sperm count’) has declined by more than 50% among ‘Western’ men between 1973 and 2011. ‘Western’ populations included North America, Europe, Australia and New Zealand. They were contrasted with ‘Other’ populations from South America, Asia, and Africa, where sperm count declines were not found to be statistically significant (Levine et al., 2017, pp. 650, 654).

In this paper, we locate the interpretation by Levine and colleagues of recent human sperm count trends within a conceptual framework that we call the Sperm Count Decline (SCD) hypothesis. We contrast the SCD hypothesis with a proposed alternative, which we call the Sperm Count Biovariability (SCB) hypothesis. SCD and SCB represent two interpretive frameworks for the same data and generate different research programmes for the study of sperm count trends in human populations. We contrast four propositions of the SCD and SCB hypotheses, summarised in Tables 1 and 2, and argue that SCB offers a more analytically rigorous, empirically driven, and generative framework for designing data collection protocols and interpreting population trends in human sperm counts than does SCD. The SCB offers improved tools both...
for the re-analysis of existing epidemiological data and for the design, analysis, and interpretation of future research on sperm counts, environment, and men’s health.

**History and social context of the sperm count decline debate**

Levine et al. (2017) are not the first or most recent researchers to suggest that contemporary social and physical environments harm sperm production in men. Debates over whether sperm counts are declining and what such decline might mean have been taking place for decades (Carlsen et al., 1992; Daniels, 2006; Nelson & Bunge, 1974; Sengupta, Dutta et al., 2017; Sengupta, Nwagha et al., 2017; Sengupta, Borges et al., 2018; Sengupta, Dutta et al., 2018, Swan & Colino, 2021). Recent research on sperm count decline is associated with a broader research programme postulating a rise in andrological pathologies, resulting from a hypothesised Testicular Dysgenesis Syndrome, or TDS (Akre & Richiardi, 2009; Skakkebaek et al., 2001). Here we focus our analysis on Levine and colleagues’ meta-analysis for three reasons.

First, the scientific impact of Levine et al. (2017) is significant: their paper was the top cited paper in 2017 in *Human Reproduction Update*, a leading journal in the field (Human Reproduction Update, Oxford Academic, n.d.; Web of Science, 2020).

Second, the paper avoids some of the methodological weaknesses that provoked scepticism within the scientific community about other claims of sperm decline (Carlsen et al., 1992; Daniels, 2006), granting their findings unprecedented authority among reproductive scientists (Science Media Centre, 2017). Specifically, they created rigorous criteria so as to only include studies that used commensurate sperm count measurement techniques; attempted to extract data on a wide range of potential confounders; implemented a systematic quality control protocol; conducted several sensitivity analyses; and excluded samples that were selected for men with infertility or for the re-analysis of existing epidemiological data and for the design, analysis, and interpretation of future research on sperm counts, environment, and men’s health.

The Levine et al. (2017) meta-analysis treats ‘Western’ and ‘Other’ as appropriate, implying that there are stable differences in bodies and environments across these categories and across time. For
example, it assumes that it is conceptually warranted to compare average population sperm counts among ‘Western’ men in the 1970s to ‘Western’ men in 2013. Further, it implies that the sperm counts of 1970s Western men represent an optimum compared to 2013 average sperm counts. This formulation situates men’s bodies and environments labelled ‘Western’ as exemplary, natural, and now imperilled.

Men’s Rights Activist (MRA) groups, an outgrowth of the men’s liberation movement formed in reaction to feminism and a perceived crisis of masculinity in contemporary societies (Manne, 2018; Stern, 2019), seized on the interpretation of Levine et al. (2017) that contemporary chemical and social environments are hostile to Western masculinity. Discussions on some of the most prominent MRA online forums have framed the study as proof that Western civilisation has doomed itself (T. Moore, 2018). Far right conspiracy theorist Alex Jones linked the drop in Western men’s sperm counts to the eroding social status and feminisation of men, and connected it to claims of declining testosterone levels and derisive ‘soy boy’ theories about the feminising effects of oestrogen-containing soy products (Henderson, 2018). Levine et al. (2017) appeared to affirm the idea that men’s putatively declining social status was indeed behind the declines: ‘Social factors could definitely influence this … We are animals. The social rank, the socioeconomic position, is important,’ Levine was quoted as saying in The New York Times (Bowles, 2018).

**Contrasting the sperm count decline and sperm count biovariability hypotheses**

In this paper we contrast the Sperm Count Decline and Sperm Count Biovariability hypotheses. We understand a hypothesis as not only a set of propositions open to empirical testing, but also as a set of implicit and explicit model-theoretic assumptions about the world that provides a framework for collecting and interpreting new and existing data and setting research agendas.

The SCD hypothesis interprets data on sperm count over time as a metric of men’s potential fertility, a proxy for men’s health, and an assay of environmental quality. According to SCD, a decline in ‘Western’ sperm counts from 1970s levels indicates a decline in male fertility, health, and a sign of a degrading environment. By contrast, the SCB hypothesis allows for the possibility of both pathological and non-pathological variation in sperm counts across populations and time. SCB begins with the premise that, above the threshold necessary for fertility, there is no basis to assume that high average population sperm counts are optimal. Nor is there any reason to believe that sperm counts in the 1970s are a species-typical baseline. SCB posits that sperm count varies within and across bodies in ways that are compatible with health such that a decline in an individual or population may not necessarily signal danger to fertility or well-being. We emphasise that while SCB invites a wider explanation and interpretation of sperm count trends, it does not exclude the possibility of sperm count decline or that decline may carry implications for men’s health and fertility. The SCB hypothesis provides a framework for exploring the trends identified by Levine et al. (2017) that considers the possibility that these trends can be explained by benign or adaptive variation in sperm counts in relation to diverse contexts and factors. Rather than treat nations or regions of global wealth as proxies for stable populations or biologically meaningful environments, SCB calls for testing links

**Table 2.** Contrasting interpretations of sperm count trends: Sperm Count Decline (SCD) vs. Sperm Count Biovariability (SCB) hypotheses.

<table>
<thead>
<tr>
<th>Sperm Count Decline (SCD)</th>
<th>Sperm Count Biovariability (SCB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Sperm and Fertility</td>
<td>Male infertility is a complex biological and social phenomenon that is not predicted by the single metric of sperm count.</td>
</tr>
<tr>
<td>Fertility scales proportionately with sperm count, and a recent drop in sperm count means that fertility is already imperilled.</td>
<td></td>
</tr>
<tr>
<td>(2) Sperm and Overall Health</td>
<td>The connection between sperm count and health is mediated by the individual’s recent experience and prior life history. The short time scales for sperm regeneration make sperm non-ideal for tracking health status.</td>
</tr>
<tr>
<td>Average population sperm count is a biomarker of the overall health status of men.</td>
<td></td>
</tr>
<tr>
<td>(3) Sperm Count and Environmental Pollutants</td>
<td>The relationship between endocrine disrupting chemicals, geography, and sperm count has not been established by prospective studies across populations that control for confounders and use repeat individual measures in combination with a wide variety of social and biological measures.</td>
</tr>
<tr>
<td>Declining sperm counts since 1973 in ‘Western’ countries are likely the result of environmental endocrine disrupting chemicals.</td>
<td></td>
</tr>
<tr>
<td>(4) ‘West,’ ‘Other,’ and Nations as Units of Population</td>
<td>Sperm counts rise and fall and vary across populations and time, and within individuals, in a dynamic, non-linear and non-continuous fashion.</td>
</tr>
<tr>
<td>The average sperm count in a population sample and in a given year is sufficiently meaningful for comparing populations and describing population trends. Global sperm decline is linear and continuous.</td>
<td></td>
</tr>
</tbody>
</table>
between specific developmental and proximate stimuli and sperm count outcomes, recognising human biological variation as local and situated (Lock, 2017; Niewohner & Lock, 2018).

From an SCB perspective, the data points that make up the 2017 meta-analysis simply demonstrate that sperm count varies across bodies, ecologies, and time periods. Examining the same data and background literature with a different set of assumptions, SCB argues that the interpretation that population sperm counts vary within a wide optimum with little consequence for fertility is at least as plausible as the interpretation that steady decline occurs.

We argue in favour of the SCB as a framework for interpreting population trends in human sperm counts. It identifies testable hypotheses that include both pathological and non-pathological explanations for and outcomes of observed variation in sperm counts. Table 2 contrasts the propositions of the SCD and SCB hypotheses. In the following sections, we analyse each proposition pair in turn.

1. Sperm count and men’s fertility

The SCD hypothesis contends that lower average population sperm counts portend higher rates of male infertility, positioning sperm count decline as a marker or cause of reproductive crisis for the human species. Levine et al. (2017) for example, infer that ‘declining mean [sperm count] implies that an increasing proportion of men have sperm counts below any given threshold for sub-fertility or infertility’. Levine et al. (2017) link this to claims of increasing ‘economic and societal burden of male infertility’ (p. 649).

There is little evidence that this is true. Levine et al. (2017) contend that the high circa 1973 numbers represent normal, healthy, and natural levels, while today’s numbers represent a crisis and decline from a prior optimum. But current Western average sperm counts reported by Levine et al. (2017) for men unselected for fertility are well within the ‘normal’ range, defined by the World Health Organisation (WHO) as 15–259 million per mL for individuals (World Health Organization, 2010, p. 224). That is, the Levine et al. (2017) study reports a population average decline from ‘normal’ (99 million sperm per ml) to ‘normal’ (47 million sperm per ml). Furthermore, in absolute numbers, the 2009–2011 Unselected Western sperm counts (47.1 million/mL), which are ostensibly cause for alarm, are in fact relatively close to absolute sperm counts in ‘Other’ countries back in the 1978–1983 period (66.4 million/mL for Fertile Other, 72.7 million/mL for Unselected Other) and in the 2010–2011 period (75.7 million/mL for Fertile Other, 62.6 million/mL for Unselected Other).

Male infertility is a complex biological and social phenomenon that cannot be understood in terms of the single metric of sperm count (Guzick et al., 2001). Though azoospermia (sperm count of zero) guarantees infertility, researchers have found that some men with low sperm counts can conceive, while others with higher counts cannot (Patel et al., 2018; Wang & Swerdloff, 2014). Guzick et al. (2001) demonstrate that even sperm concentrations in the so-called sub-fertile range of less than 13.5 million/mL ‘do not exclude the possibility of normal fertility’ (p. 1392). Of note, the 2010 WHO reference values for semen parameters do not predict infertility, as the values were determined by studying fertile men; therefore, while the top 95% of sperm concentrations in the sample were taken to be the reference range, all of the men with sperm concentrations below the 5th centile were also fertile (Chiles & Schiegel, 2015; Cooper et al., 2010). Other studies from across the world have similarly confirmed the fertility of men below the WHO reference values (Haugen et al., 2006; Tang et al., 2015; Zedan et al., 2018).

Clinicians do not report proportionate increases in infertile men presenting for clinical consultation over Levine et al.’s study period (Inhorn & Patrizio, 2015). As urologist Peter Schlegel remarked in The New York Times in reference to the Levine et al. (2017) meta-analysis, ‘If you had a decrease in sperm count in the 50 to 60 percent range, we would expect the proportion of men with severe male infertility to be going up astronomically. And we don’t see that’ (Bowles, 2018). There is insufficient evidence to support claims of increasing rates of male subfertility in recent decades (Inhorn & Patrizio, 2015).

We note that there exists no species optimum in many other measures of reproductive function in men and women. As a concrete example, the gonadal steroid hormones testosterone, oestradiol, and progesterone are necessary to fertility (Dohle et al., 2003; Laufer et al., 1982; Welt et al., 2003). Researchers have documented significant variation in these hormones within and across populations and within individuals over time (Bjørnerem et al., 2006; Stanton et al., 2011; Vitzthum et al., 2004). As the WHO does with sperm count, researchers validate and publish non-pathological ranges for gonadal steroid hormones for use in clinical evaluation (Bhasin et al., 2011; Elmlinger et al., 2002). Within those ranges, higher levels are not
considered absolute signals of better fertility or health (Bribiescas, 2016).

2. Sperm count as an assay of men’s health

SCD interprets sperm count decline as a biomarker of declining overall health status among men. Citing studies associating reduced sperm count and ‘increased all-cause mortality and morbidity’ (Levine et al., 2017, pp. 647, 649, 654), Levine et al. (2017) hypothesise that average population declines in sperm counts represent a ‘canary in the coal mine’ for male health across the lifespan (p. 654; see also Skakkebaek et al., 2001). This metaphor suggests that low sperm counts are not only a barometer of men’s current health, but also a warning sign of future risks.

While there is evidence of a relationship between abnormal semen parameters and poor health status (Eisenberg et al., 2014), there is little evidence that average sperm count by itself is a valid summary measure of health status of men within a population. Recent work in a population in Córdoba, Argentina, suggests that, while semen parameters decline with age, lifestyle and health factors such as obesity, alcohol, and smoking have only modest associations with decline (Veron et al., 2018). Similar findings exist with respect to other semen parameters: a 1999 study of 939 UK men found no relation between sperm motility and common lifestyle factors such as consumption of alcohol, use of tobacco or recreational drugs, or high body mass index (Povey et al., 2012).

Specific relationships between sperm parameters and developmental and current conditions, including health status, remain to be established. Sperm variability can reflect endogenous and exogenous stimuli on both short and longer time scales. Spermatogenesis is a 42–76 day process (Misell et al., 2006). Interventions can occur at any point from the first division of the spermatogonia to the mature sperm’s journey through the epididymis (Chenoweth & Lorton, 2014). Research on livestock indicates that, depending on the developmental stage of their influence, effects can be permanent or may resolve. For example, enhanced nutrition in early life increased adult sperm production in bulls, but later-life nutrition could not compensate for early-life nutritional deficits (Kastelic, 2013). Seasonal climate variation, however, had only a transient effect on sperm parameters in bulls (Valeanu et al., 2015). Further research is needed to establish whether the same range of developmental and transient effects can be found in humans.

Prospective study models that use repeat individual measures in combination with a wide variety of social and biological measures are needed to identify potential confounders and causal variables in sperm biovariation. Such variables include transient exposures such as heat or tight clothing; the stimulus conditions under which the sample was collected, including available arousal material and duration of arousal pre-ejaculation; lifestyle factors including activity and diet; and developmental or environmental exposures like maternal smoking, pollutants, and endocrine disruptors (for examples of existing cross-sectional studies along these lines, see Gaskins et al., 2015; Inhorn et al., 2008; Priskorn et al., 2018). Without longitudinal individual and population data with sufficient ecological granularity, causal claims about the relationship between average population sperm counts and environments or lifestyles cannot be empirically substantiated.

In any case, the connection between sperm count and health is mediated by the individual’s recent experience and prior life history. For example, increased exercise does not have a stable relationship to sperm production, in part because the effect of exercise is mediated by current fitness level (Ibáñez-Perez et al., 2019; Józków & Rossato, 2017; Rosety et al., 2017). Factors such as seasonal temperature and illness do not have uniform effects on the sperm production process for similar reasons. Given the range of relationships between stimulus and effect in spermatogenesis, sperm count is not an independent metric of human well-being.

3. Sperm count and environmental pollutants

In line with the TDS hypothesis (Bay et al., 2006; Skakkebaek, 2016), SCD asserts that the likely causes for sperm count declines among ‘Western’ populations are endocrine disruptors and other environmental pollutants introduced by industrialisation, as well as changes in men’s lifestyles. Levine et al. (2017) write that, ‘sperm count and other semen parameters have been plausibly associated with multiple environmental influences, including endocrine disrupting chemicals, pesticides, heat and lifestyle factors, including diet, stress, smoking and BMI’ (Levine et al., 2017, p. 649). In a Guardian article titled, ‘Sperm counts are on the decline - could plastics be to blame?’, Levine identifies endocrine-disrupting chemicals (EDCs) such as plastics as a major cause of dropping sperm counts (Carr, 2019).
While environmental context undoubtedly affects men’s health, empirical research to date does not support a stable causal relationship between EDCs – exogenous chemicals that interfere with hormone action, typically through mimicking endogenous hormones and binding to protein receptors – and any indices of sperm health, including sperm count, sperm motility, and fertility (Bonde et al., 2016; Zamkowska et al., 2018). Scientists have approached questions about the impact of EDCs on reproductive function through animal models and human studies. In animal studies, male rodents are exposed to specific quantities of EDCs in a controlled environment, and systematically examined for effects on their reproductive health. In contrast, human clinical and epidemiological studies are primarily observational, studying the sperm of human males in the general population who were accidentally exposed to unspecified levels of EDCs.

The strongest evidence for the impact of EDCs on human populations lies in their action as somatic carcinogens (Soto & Sonnenschein, 2010). Although reproductive cancers could plausibly lower sperm count, this pathway cannot explain the patterns reported by Levine et al. (2017) as they exclude cancer patients from their study. EDC exposure is also associated with risk for a wide range of health conditions outside of its effects on reproductive health, including non-reproductive cancers, diabetes, thyroid disorders, and neurological conditions (Gore et al., 2015). Some scientific research suggests that EDCs can have reproductive, neurological, and immunological effects on developing human foetuses (of both males and females), but more research is required to establish the exact relationship (Abaci et al., 2009; Bonde et al., 2016).

Even if EDCs cause a decline in sperm count, higher levels of industrial pollutant exposure in the West cannot explain the divergent trends in Levine et al.’s categories of ‘West’ versus ‘Other.’ Scientists have used the global distribution of plastics as a geological indicator of the extent of human altered landscapes (Zalasiewicz et al., 2016). It is widely established that the inequities of global capitalism disproportionately burden the global poor and indigenous peoples with the consequences of toxic pollution (Martinez-Alier et al., 2016). Substantial evidence suggests that pesticide poisoning is an equal or greater problem in low- and middle-income countries as in high-income countries (Jørs et al., 2018); the World Health Organization (2016) reports that 98 percent of people in urban low- and middle-income countries are exposed to unhealthy levels of toxic pollution.

The study period of 1973 to 2011 included in Levine et al. (2017) accompanied increasing global levels of industrial pollution (He et al., 2002; Karan & Bladen, 1976; Ramakrishnan, 2018). As a detailed example, consider two studies from Hyderabad, Andhra Pradesh, both included in Levine et al. (2017): one from a lead-acid battery manufacturing facility in Patancheru District and another at an anonymous lead welding facility (Danadevi et al., 2003; Vani et al., 2012). The 1970s initiated a rapid intensification of environmental pollution, in the form of waste disposal, air pollution, wastewater effluents, and occupational exposures in this region (Danadevi et al., 2003; Vani et al., 2012). Half of India’s rivers today are polluted by industrial wastewater effluents introduced in the 1970s; in Hyderabad, water sources were contaminated with industrial metals as early as 1983 (Karan & Bladen, 1976; Prahalad & Seenayya, 1988). Vegetables grown and consumed in urban Hyderabadi areas are packed with lead, cadmium and chromium, and bodies of water sampled are saturated with EDCs (Kiran Kumar et al., 2016; Ramakrishnan, 2018; Srikanth & Papi Reddy, 1991). In summary, evidence does not support the claim that increased exposure to environmental pollutants in the nations categorised by Levine et al. (2017) as ‘Western’ could be a plausible driver of distinctions between average population sperm count in ‘Western’ compared to ‘Other’ nations.

### Table 3. Sperm concentration in the first and last years of the Levine et al. (2017) meta-regression analysis, for all men and by fertility and geographic groups ‘Western’ and ‘Other.’

<table>
<thead>
<tr>
<th>Category</th>
<th>N</th>
<th>First year</th>
<th>First year sperm count (millions/mL)</th>
<th>Last year</th>
<th>Last year sperm count (millions/mL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All men</td>
<td>244</td>
<td>1973</td>
<td>92.8</td>
<td>2011</td>
<td>66.4</td>
</tr>
<tr>
<td>Unselected Western</td>
<td>110</td>
<td>1973</td>
<td>99.0</td>
<td>2011</td>
<td>47.1</td>
</tr>
<tr>
<td>Fertile Western</td>
<td>65</td>
<td>1977</td>
<td>83.8</td>
<td>2009</td>
<td>62.0</td>
</tr>
<tr>
<td>Unselected Other</td>
<td>30</td>
<td>1986</td>
<td>72.7</td>
<td>2010</td>
<td>62.6</td>
</tr>
<tr>
<td>Fertile Other</td>
<td>39</td>
<td>1978</td>
<td>66.4</td>
<td>2011</td>
<td>75.7</td>
</tr>
</tbody>
</table>

4. ‘West,’ ‘Other,’ and nations as units of population

Levine et al. (2017) extracted 244 estimates of average sperm counts from human sperm samples collected over the period 1973–2011 and reported in English-language publications, representing 42,935 individual men across the globe. They report that the average sperm concentration across Unselected ‘Western’
populations was 99 million/ml in 1973, declining to 47.1 million/ml in 2011, or 52.4% overall (Table 3). Sperm declines were only statistically significant in studies in ‘Western’ countries (Levine et al., 2017, p. 654), while [n]o significant trends in SC [sperm concentration] or TSC [total sperm count] were seen in ‘Other’ countries overall, or for Unselected or Fertile men separately (Levine et al., 2017, p. 652).

The study design of Levine et al. (2017) separated men along axes of fertility and geography. First, they categorised men as ‘Unselected,’ meaning that it was not known whether or not they had conceived a pregnancy, or as ‘Fertile,’ meaning that they had conceived a pregnancy. They next disaggregated men by the nation of the study in which they participated: Europe/Australia; North America; and ‘Other,’ which included South America, Asia, and Africa.

Constituting ‘Western’ and ‘Other’
Notably, the model used by Levine et al. (2017) generated statistically significant declines in sperm concentration over time for both Unselected and Fertile Europe/Australia cohorts, and for the Unselected North America Cohort, but not for the Fertile North America cohort (p = 0.29) or either of the Other cohorts (p = 0.30 for Unselected Other, 0.41 for Fertile Other) (Levine et al. (2017) Table S3). In the final published study, Levine et al. (2017) aggregated North American and Europe/Australia data to create a ‘Western’ cohort. Their justification was the similarity of effect magnitude in the data (despite one subgroup – Fertile North America – being statistically insignificant) and that North American data comprised only 16% of the estimates. In the final model, both Unselected and Fertile Western had statistically significant negative effects. In other words, sperm count declines in North America among Fertile men, which were not previously significant (p = 0.29), gained manufactured significance (p = 0.033) by being weighted with the European/Australian data in the final model.

It is justifiable to explore multiple aggregations of data along hypothesis-driven inquiries. However, the reframing of a statistically insignificant decline in fertility among Fertile North American men implies a level of certainty that the data do not support. When this certainty is adopted by public-facing reporting, it not
only contributes to unfounded panic over ‘Western’ fertility, but also may influence the course of future research programs.

The data included in the meta-analysis are sparse by any measure. Global coverage is mottled and asymmetric. Levine et al. (2017) recognise that data on sperm count in ‘Other’ countries is much sparser than in ‘Western’ countries, as illustrated by Figure 1 and Table 3. Less evident yet still important is the quantitative and qualitative variation in the data points at the level of the nation and the region. For example, the preponderance of sperm count studies over a range of time periods from several major Danish cities included in the meta-analysis might allow researchers to describe general population trends in sperm count within Denmark. However, the same number of studies, or fewer, conducted at disparate times in such large and heterogeneous countries such as India or China cannot hope to capture the same granularity of data by averaging sperm counts.

Studies of men unselected by fertility (i.e. men assumed to be representative of the general population in a given geographic area) included in the meta-analysis vary in study design and sample composition across geographic location. For example, control groups in studies of the impact of an environmental exposure on sperm quality were extracted for inclusion in Levine et al. (2017). Studies conducted in this way contribute more samples to some national data pools than others. For example, of 10 separate studies conducted in Denmark, four (40%) were interested in the impact of a specific exposure (e.g. pesticides or maternal folic acid) on sperm quality. By contrast, 13 of the 16 studies in the United States (81%) are exposure studies, looking at the effects of a chemical exposure, smoking, stress, or a medical condition such as cryptorchidism on sperm quality. And 100% of the five studies used for Unselected samples in all of Central and South America were designed to study sperm quality in the context of a specific exposure, whether pesticides, contaminants, or a medical condition. This is important because the controls in these exposure studies are often convenience samples relative to study subjects. For example, a study in Mexico City on rubber factory workers exposed to hydrocarbons used a control group consisting of employees working in the factory’s administrative offices (De Celis et al., 2000), and a study in San Francisco on sperm quality in anaesthesiologists had anaesthesia residents serve as controls (Wyrobek et al., 1981). As Fleiss and Gross (1991) explain, factors other than exposure may affect whether a sample is a case or a control, and these confounding variables can obscure the associations of interest.

**Interpreting average sperm count in a nation**

The SCD treats nations and continents as bounded populations, with men unselected by fertility described as ‘more likely to be representative of the general population’ in that nation or continent (Levine et al., (2017), p. 655). That is, continents or nations are conceptualised as population samples that can be used to compare, for example, average 1973 sperm counts to average 2011 sperm counts.

Within this paradigm, the categories ‘West’ and ‘Other’ rely on a particular vision of a static national population that obfuscates the role of several types of migration in continually redefining how these populations are constructed. Since the 1970s, repeated, large-scale, highly varied movements of populations have occurred across national borders, and in particular between nations categorised by Levine et al. (2017) as ‘Other’ to ‘Western’ countries. Yet these movements, from East and South to North and West, from rural to urban, and across many kinds of differently polluted and polluting ecologies, are lost entirely in the racial/national geopolitical categories of difference uncritically embraced by current instantiations of the SCD hypothesis. If biological variables such as sperm count are to be understood as ecologically dependent at the population level, patterns of migration since the 1970s that have fundamentally reshaped nations categorised as ‘West’ and ‘Other’ must be taken into consideration.

During the time period covered by the Levine et al. (2017) meta-analysis, patterns of migration have redistributed formerly concentrated populations into a contingent of increasingly heterogeneous cities and states, predominantly in the Global North and West. In Western Europe, decolonisation as well as the proliferation of guest worker programmes to meet the needs of a broadly booming post-war economy brought individuals from former colonies distributed over three continents, combined with workers from North Africa and Southern Europe, into Northern and Western Europe (Moch, 2003; Van Mol & de Valk, 2016). Sweden, for instance, which historically had higher emigration than immigration, saw a rapid population change after World War II. Immigration rates peaked in 2013. As a result, 24% of the population is now foreign-born, and its ethnic composition has also shifted. Migration from Africa again rose in the 1990s and migration from across Asia and Latin America into Western Europe rose significantly from the turn of the 21st century (Pellegrino, 2004; Van Mol & de Valk,
decades of cyclical immigration from rural areas towards urban industrialisation, emigration abroad both to ‘West’ and ‘Other’ nations, and a recent reversal of this process wherein Western-trained Indians return as Hyderabad grows in its transnational, globally connected contemporary networks (Chacko, 2007). The social and environmental experience of growing up and living in South India in the 1970s, for example, is not comparable to that of South India in the 2010s. It is unclear how Levine et al. (2017) locate ‘India’ in their analysis, whether as a place with a set of defined ecological conditions, as a group of people, or as a place that might have changed over time ecologically but where the population has remained constant enough to allow for disambiguation of the effects of place on sperm production from any other effects. We suggest that it is not obvious that the genetic composition of a population in a given place remains the same over time. Nor is it certain that the people in a given place have experienced the developmental environment of that place, or that the place has remained ecologically stable (or not) in predictable or documented ways.

A biovariability framework emphasises that the appropriate unit of analysis to understand relationships between ecology and average population sperm count is a spatiotemporally continuous population, in which bio-environmental and socio-cultural exposures and their outcomes can be tracked over time within and among individuals at multiple time points. Levine et al. (2017) assume that geographic regions such as nations and continents over the period from 1973 to the present day represent such populations. Although it may sometimes be the case that geopolitical boundaries can meet these criteria, the highly dynamic history of migration and environmental change since 1973, wrought by increasingly globalised processes, indicates that the nation-level sperm count averages utilised by Levine et al. (2017) are inappropriate categories for understanding sperm count epidemiology.

Conclusion: A biovariability framing for research on the future of sperm

For half a century, scientists have worried about the possible decline of human sperm counts. In 1974, Nelson and Bunge reported in the journal Fertility and Sterility that ‘something has altered the fertile male population to depress the semen analysis remarkably’ (Nelson and Bunge, 1974, 507). Their result was confirmed by follow-up studies, each reporting drops in sperm counts and eliciting significant public attention
(see Auger et al., 1995; Carlsen et al., 1992; Irvine et al., 1996; Osser et al., 1984).

These studies were met with considerable scepticism. Critics pointed out that studies claiming sperm count declines contained methodological flaws, ranging from issues with sample selection to the use of incommensurate counting techniques, failures to account for confounds, and misleading statistical models (see Bahadur et al., 1996; Fisch, 2008; Merzenich et al., 2010; Pacey, 2013).

Furthermore, other studies reported contradictory findings: that sperm counts were not in fact declining, at least not everywhere; in some locations, sperm counts were even found to be increasing (see Becker & Berhane, 1997; Fisch et al., 1996; Jørgensen et al., 2012).

The meta-analysis by Levine et al. (2017) of global sperm count change over time sought to avoid the methodological pitfalls of the studies that came before. Its findings received wide uptake across varied arenas, including by andrologists, advocates of the endocrine disruptor hypothesis, environmentalists, men’s rights/alt-right activists, and popular media. Public translations of sperm decline claims often apply evidence of population-level global trends to men’s individual lives and encourage men to make decisions about their health and fertility based on these claims. In the media, the results have been framed in terms of a crisis for Western men and masculinity (e.g. Bowles, 2018; Corbyn 2021; Davis, 2017; Halpern, 2018; McKie, 2017; Salam, 2017; Whitworth, 2021). Authors of the 2017 paper have contributed to these narratives, including in a new book Count Down (Swan & Colino, 2021), which claims that men face ‘environmental emasculation’ (p. 2) and that endocrine disrupting chemicals not only threaten sperm count but ‘blur’ gender differences in ‘language-development… and many other qualities’, potentially resulting in ‘gender dysphoria’ (p. 59–60).

Choice of interpretive framework for approaching human sperm count studies matters. Infertility is a deeply personal and significant health experience for many people. Scientific claims about sperm count decline are also entangled with cultural anxieties about the perceived decline of masculinity (Daniels, 2006; Inhorn, 2004). The ‘silent shame of male infertility’ (Oaklander, 2019) leaves many men feeling a ‘failure at fatherhood’ (Anthony, 2018) and a ‘threatened sense of masculinity’ (Sylvest et al., 2018). As social scientists have shown, representations of sperm are ‘deeply contextually situated within a crisis of masculinities’ (Almeling, 2020; T. Moore, 2018, p. 73). In discourse analyses of media coverage of male infertility, men are represented as ‘vulnerable and threatened by forces outside their control’ (Gannon et al., 2004). Additionally, scientific claims surrounding threats to masculinity and fertility in a comparative geographic, national, or ethnic context ignite and inflame powerfully divisive discourses concerning racial, gender, and national futures (Stern, 2019; Wahlberg, 2018).

One could argue that the Sperm Count Decline hypothesis is analytically separable from the racial/ethnic, continental, and national interpretive framework invoked by Levine et al. (2017). This, however, is not how the hypothesis operates in the analysis by Levine et al. (2017) or in the paper’s reception. Hypotheses are not merely sets of propositions but socially-situated bundles of implicit and explicit assumptions. Given the data that we have on sperm count averages over time – and the Levine et al. (2017) meta-analysis is the most complete dataset available – one cannot arrive at a finding of global sperm count decline without accepting ‘West’ and ‘Other’ as geospatially meaningful categories for aggregating men’s reproductive histories and potentialities.

A biovariability framework is better poised to test central propositions of the decline hypothesis and to interrogate – as well as investigate explanatory frameworks beyond – binary categories such as ‘West’ and ‘Other.’ Through this lens, only geospatial categories that attend to local environmental context are appropriate for data aggregation (Lock, 2017; Niewohner & Lock, 2018). Systematically grouping studies in a multitude of ways that account for local social, political, and environmental context, recording significant and non-significant results, and then reporting the totality of those findings would offer a more rigorous understanding of the contexts and causes of global sperm count biovariation. The biovariability hypothesis remains open to new explanatory categories attentive to the ecological changes that result from social and historical context.

Researchers must take care to weigh hypotheses against alternatives and consider the language and narrative frames in which they present their work. In addition to its explanatory virtues, we argue that biovariability offers a more promising framework than does ‘sperm decline’ for attending to these imperatives. The SCD hypothesis produces a picture of crisis around declining sperm counts among Western men, but treats the already lower average sperm counts of non-Western ‘Other’ men as outside of the umbrella of concern and crisis. With the focus on decline in
‘Western’ countries, the discourse is often implicitly (or explicitly) racialised, implying imperilled white male fertility. As a result, the SCD is presently being deployed by forceful and rising rightwing nationalist movements around the world to support inflammatory assertions about the decline of the ‘West,’ the feminisation of males, and the differential fertility rates of people of colour (e.g. Moore, 2018).

The claim of significant trends in sperm count decline only in countries described as ‘Western,’ and not in those described as ‘Other,’ condenses our concerns about the SCD hypothesis and its implicit assumptions. A biovariability approach posits that the appropriate unit of analysis for relating average population sperm counts to ecology and health is a spatio-temporally continuous population in which exposures (e.g. pollution, poverty) and outcomes (e.g. health metrics) can be tracked over time. In other words, this approach involves comparing sperm count averages across populations by sampling the same men at multiple time points. It should not be assumed at the outset that nation/state boundaries demarcate such populations; in many cases they do not, although some may. Aggregate analyses of populations should be based on explicit, empirically supported premises about shared characteristics as well as well-motivated hypotheses about the causes of sperm variation (e.g. BMI and smoking).

The SCD hypothesis contends that sperm count is declining and that sperm counts will likely continue to decline at similar rates without interventions; it asserts that lower sperm counts reflect declining health status among men, a result of exposure to environmental toxins as well as changes in men’s lifestyles; and, it claims that as average sperm counts grow lower, human male fertility is in peril. We propose an alternate perspective, the SCB hypothesis. According to this view, sperm count likely varies across bodies within a wide range, much of which can be considered normal from the perspective of reproductive function and well-being, or even optimal, in particular adaptive contexts. We stress that the SCB hypothesis does not rule out the possibility that the average sperm count of a well-defined population could decline due to negative environmental exposures, or that this may carry implications for men’s health and fertility. However, claims of this sort require evidence of a causal relationship between sperm count and life historical and ecological factors, which should be tested against the possibility that a wide range of non-pathological variation is typical for sperm count.

Acknowledgements

Thank you to Hagai Levine for sharing the full dataset from Levine et al. (2017) for our reanalysis, and to Jeff Blossom and Samantha Carr of the Center for Geographic Analysis at Harvard University for producing the maps.

Disclosure statement

The authors report no conflict of interest.

References


Bjørnerem, A., Straume, B., Oian, P., & Berntsen, G. K. (2006). Seasonal variation of estradiol, follicle stimulating hormone, and dehydroepiandrosterone sulfate in women and


Elmlinger, M. W., Kuhnle, W., & Ranke, M. B. (2002). Reference ranges for serum concentrations of lutropin (LH), follitropin (FSH), estradiol (E2), prolactin, progesterone, somatomedins, growth hormone, thyrotropin (TSH), sex hormone-binding globulin (SHBG), and dehydroepiandrosterone sulfate (DHEAS). Endocrine Reviews, 23(6), 1169–1186. https://doi.org/10.1210/jcem.2002.02002


