Global decline in semen quality: ignoring the developing world introduces selection bias

Raywat Deonandan
Marya Jaleel
Interdisciplinary School of Health Sciences, University of Ottawa, Ottawa, Ontario, Canada

Abstract: Multiple studies from around the world have suggested that semen quality is declining globally. However, all studies suffer from variable semen sampling criteria, selection bias with respect to the types of men volunteering to participate, and a bias with respect to a tendency to examine only samples from high-income countries. This heterogeneity in approaches, especially given the undersampling of rural and less affluent men from low-income countries, calls into question researchers’ claims of universally declining semen norms.

Keywords: human semen, quality, global, developing world

Introduction

For decades, scientists have collated data from the collection of semen samples from the world over to establish norms for use by the wider scientific community. Some standardization in the collection process has been guided by the World Health Organization’s various published guidelines for semen management and resulting expectations of quality. The World Health Organization manuals for the examination of human semen1–4 are widely used as a source of standard methodology for both population and laboratory studies. While some studies on single populations show no change over decades,5,6 long-term analyses suggest that, overall, semen quality is declining.7–14

This downward trend was most famously declared to be worldwide by Carlsen et al8 who found a decline in sperm quality after analyzing published studies from 1938 to 1991. Their paper led to much commentary in the literature, most of it focusing on possible causes of the decline, ie, environmental exposure, nutritional, behavioral, or genetic factors. Possible causes discussed in the lay media tend to fixate about something called the “estrogen hypothesis”,15 which points the finger at variable levels of environmental estrogen. Blame has variably been placed on a host of influences, such as societal overuse of soy products, which may act as an environmental phytoestrogen; overuse of the female contraceptive pill, which may result in estrogens being excreted in cities’ water supplies; or the disposal of some industrial products, which may also create an unusual hormonal environment; or some other as yet unidentified environmental factor.16 Proposed behavioral factors include a rise in obesity, diabetes, and overall poor health and unfitness in men,17 or factors lumped under the umbrella term “urban lifestyle”.18 There is some evidence that this decline is associated with an increase in other impairments of the male reproductive system19 and even with mobile phone usage.20 It also appears as if this is a uniquely human trend, because analysis of other mammalian species does not show comparable
declines in semen quality. More nuanced discussions of possible causes of semen decline involve complex biological-environmental interactions.

There are some dissenting studies, usually limited to restricted geographies. Quite compelling is a recent prospective Danish study suggesting that the findings of Carlsen et al are the result of a kind of detection bias, and that improved monitoring renders a finding of reducing sperm quality to be an artifact. Nonetheless, most publication activity has confirmed the trend of declining male fertility. One of the problems with such observations, as has been widely observed, is that the samples from which these trend data are derived vary greatly in terms of population characteristics and sampling criteria. What is clear is that geographical variability likely plays an important role in assessing trends in semen quality, because trend experiences in given cities seem to vary.

An important step in quality control, going forward, is to better standardize all criteria to reduce the possibility that trend findings are artifactual. Additionally, there are almost certainly differences in characteristics (behavioral, demographic, and quite possibly physiological) between men willing to give samples and those who are unwilling.

One strategy for indirectly addressing this bias is to compare samples from initial responders to sperm donation advertisements and those subsequently recruited; at least one study implies that there are significant differences in semen characteristics between these two groups, suggesting an inescapable bias inherent in all semen quality population studies. As one researcher put it, the major methodological issue here is “the use of data collected in different countries, at different times, on different populations and with different methods of semen analysis”.

While Carlsen et al famously observed the trend of declining semen quality to be worldwide, it must be pointed out that while data from developed nations are widely available, data from developing countries are sparse. This reliance on only one type of data may constitute a sort of selection bias. Low-income countries are more likely than high-income countries to be populated by men in low-technology rural environments. Such men are less likely to be exposed to industrial contaminants and are less likely to be obese. Their sexual behaviors, time spent outdoors, and clothing style (in terms of possible constriction) are also different. And while the developing world is increasingly urban, men in those cities are still less likely to be experiencing the same urban exposures of the develops world, due to differences in, eg, prevalent materials, fuels, foods, electromagnetic fields, and pathogens. This premise, that the global north and the global south are inherently different environments for male fertility, is supported by metabolic studies which suggest that the seasonal variations unique to the northern hemisphere contribute to variability in semen quality, whereas similar seasonal variations tend not to exist in the world’s poor countries.

In 2011, Zou et al studied semen samples from military personnel from scattered locations in China and found “markedly lower mean sperm concentrations, sperm counts, and sperm motility compared with WHO recommendations”, and attributed these findings to dietary, lifestyle, climate, and altitude factors. Given China’s explosive economy, demographic age profile, and the tendency for military personnel to be of higher socioeconomic status than low-income villagers, these data more reflect a developed world profile than that of lower-income nations. Indeed, the diversity within very large countries, such as China and India, presents a challenge to proper representative sampling of their resident men. Beyond their extreme domestic economic heterogeneity, such countries feature geographic extremes, ethnic diversity, class/caste distinctions, and religious and lifestyle variations that seriously complicate sampling protocols.

Further to the intracountry variability observed in large nations like China and India, an intriguing pervasive factor that may distinguish semen quality between high-income and low-income populations is diet. By comparing 30 Spanish men with poor semen quality with normospermic controls, Mendiola et al showed that frequency consumption of meaty and fatty foods, a diet characteristic of a high-income population, may negatively affect semen quality.

Another factor complicating any analysis of semen quality trends is the likely changes in legislation that happen within different nations, with respect to, among other things, donor recruitment standards. For example, Hamilton et al described the recommended changes to Britain’s donor selection criteria in 2008, reflecting attitudes regarding such factors as the size of the donor cohort and their maximum age. Similarly, laboratory quality control changes, with respect to such practices as collection and processing standards and laboratory benchmarks, may be a source of variability that confounds attempts to establish whether semen quality is really changing. The need for standardization in laboratory quality control is indicated.

There can be no doubt that sperm quality suffers from great geographic variability. The concept of a “global norm” is therefore problematic in that it necessarily ignores this underlying heterogeneity. And while some studies detect a
decline in semen quality specifically in a developing world context, the single factor that makes any such conclusion questionable is a bias in the selection of populations. The semen samples informing published analyses tend not to be random samples, but rather are typically from individuals seeking care in a reproduction clinic, or from healthy donors contributing to a sperm bank. The former are clearly at risk of being disadvantaged in some way, even when the contributor is seeking care due to a presumed female infertility issue. The second group (healthy donors) tends to be unique to the developed world, because low-income nations tend not to have well-developed public sperm donation systems. And even among the infertile, the proportion seeking care is greater in developed countries than in developing ones, further biasing samples from the latter to be from presumably more affluent individuals, who are more likely to have lifestyles and exposures concordant with those of Westerners.

Attempts by independent researchers to improve upon the history of inconstant sampling criteria fall into same the trap of oversampling developed in high-income countries. One such ambitious attempt examined semen samples from over 4500 men in 14 different countries. They were the US, China, Australia, UK, Chile, and nations from continental Western Europe. None of those countries can be considered low-income or developing nations. And if they had, those samples would have likely come from affluent, urban-dwelling individuals who are not representative of the typical man in that society.

Twice, Swan et al reproduced the meta-analysis of studies published by Carlsen et al and found that a decline was not seen in “non-Western” countries, which appear to have consisted of Nigeria, China, Egypt, Brazil, and Libya. The extent to which those countries can be considered “developing” or “low-income” is debatable. And it is likely that the semen donors in each instance were not rural, agrarian workers, but urban individuals experiencing the same environmental and occupational exposures as men from wealthier countries. Nevertheless, the authors’ finding of a difference in trends between Western and non-Western countries is intriguing, despite the extreme heterogeneity of populations and laboratory standards within their sampled studies. In opposition, an analysis of men in a supposedly nonindustrialized part of Spain showed a decrease in sperm quality over the last 30 years, lending credence to the idea that nonindustrial-related factors are involved in the observed trend.

This tendency towards selection bias is important to the formulation of theories to explain the observed global decline in semen quality, and to the development of policies and therapies to address that decline. Frankly, given the heterogeneity of sampling criteria, the tendency to sample mostly men who are seeking reproductive services, and the failure to examine samples from rural men and less affluent men in the developing world, it is incorrect to claim that declining semen quality is globally pervasive.

What is needed are new, unbiased data. Repeated prospective semen studies randomly sampled from a truly global frame, controlling for variability in affluence, age, fertility, and disease status, rural/urban lifestyle, and industrial, nutritional, electromagnetic, and ecological exposures can provide the epidemiologic resolution necessary to tease out likely causal factors from the noise, and to determine conclusively whether there exists a global crisis in male fertility. Of course, many methodological barriers would make a truly representative sampling procedure problematic, such as a difficulty in getting sufficient men in agrarian, low-income communities to participate. However, given the importance of identifying the causes and pervasiveness of observed declines in semen quality, it seems a worthy investment to consider inciting remuneration strategies to best ensure a truly representative, global sampling protocol.

Disclosure
The authors report no conflicts of interest in this work.

References


