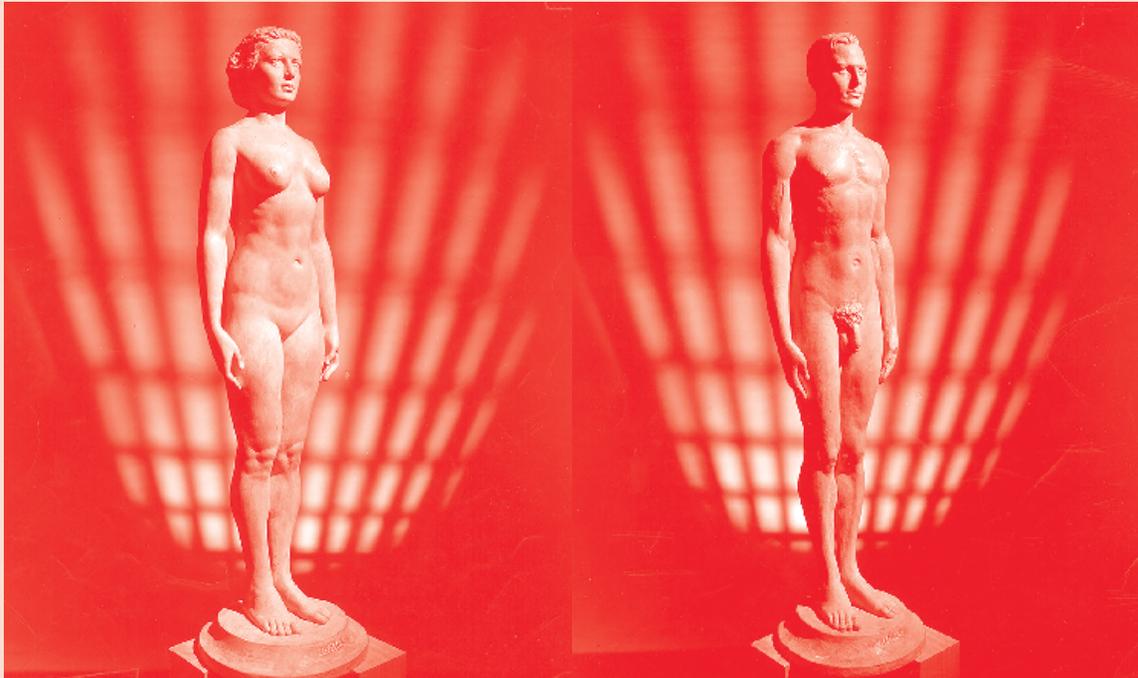


ROGER PIELKE JR. AND MADELEINE PAPE

SCIENCE, SPORT, SEX,

and the Case of Caster Semenya

Decisions about who can compete as a female athlete in world-class athletics should be informed by science, but they are ultimately subjective.



Courtesy Cleveland Museum of Natural History

In the summer of 1945, Harry Shapiro, the chair and curator of anthropology at the American Museum of Natural History in New York, revealed to the public Norma and Normman, two statues intended to epitomize the average young American male and female. These “normal” individuals were the result of years of the application of science and statistics, overseen by the New York Academy of Medicine. Norma, called “the American figure figured out by science” by William Randolph Hearst’s *American Weekly* magazine, was the product of detailed measurements of more than 15,000 women taken for use by the garment industry. Normman was the result of measurements of millions of soldiers taken during World War I.

In “Portrait of the American People,” an article announcing the project in *American History* magazine, Shapiro explained that in this case average is exceptional—even if that average defines what is normal in statistical terms. “A very fat lady and a very thin one are both rated ordinarily as less attractive than one of more average weight,” Shapiro wrote. “Obviously, then, if the average of all traits are brought together in one individual, such a person is bound to agree with the standard not only for one but for all the characters that define bodily proportion. But the combination of so many averages in one person is rare and unusual.” In taking the measurements of many individuals, the imperfections in the human form average out, and what remains is the statistically defined ideal of bodily proportion. Norma and Normman were an empirically based reflection of reality: “Norma is not meant to show what *ought* to be; she shows what *is*.” Science, Shapiro argued, reveals an underlying truth: “the average American figure approaches a kind of perfection of bodily form and proportion; the average is excessively rare.”

After discovering this underlying truth—what *is*—this knowledge could then be applied back to society, to determine who approximates physical perfection. The Cleveland Health Museum purchased Norma and Normman for an exhibit and teamed up with the *Cleveland Plain Dealer* to issue a call for applications to identify the woman who best approximated Norma. More than 3,800 women applied. The winner, a local theater attendant, only approximated Norma’s perfection, providing apparent empirical confirmation of the rarity of the perfect woman.

The racist and sexist messages accompanying Norma are now easy to spot. The statistics on which Norma was based were drawn from “college students and other thousands of native white Americans,” Shapiro explained, and Normman was the result of statistics taken from white Army soldiers. As Julian Carter of the California College of the Arts has explained, “One of the hallmarks of the ‘normal’ whiteness these statues represented was the ability to construct and teach white racial meanings *without appearing to do so*.” The statistics conveyed Norma and Normman’s version of

normal as objective, scientific, value-free. But this is a fiction: by including only “native white Americans,” the exercise excluded immigrants, men and women of color, and others from contributing to the classification of normalness.

Given this racialized approach to science, it shouldn’t be surprising that the data for Normman were collected by Charles Davenport, the founder of the Eugenics Records Office of the Carnegie Institution of Washington. Shapiro was likewise a eugenicist, who served as the president of the American Eugenics Association. Robert Latou Dickinson, the physician who oversaw the creation of the Norma and Normman sculptures along with the sculptor Abram Belskie, was another noted eugenicist. Dickinson is known for his medical sketches and sculptures related to human sexuality. Accompanying Norma and Normman in the Cleveland Health Museum were sculptures produced by Dickinson of vulvas labeled “normal,” “virgin,” “post-partum” and “lesbian.”

Far from representing what *is*, Norma was a creation of American eugenicists who wielded science to hide from view not only the actual diversity of the human form, but a deeper political agenda that today would be readily seen as racist and sexist.

The story of Norma may seem like a quaint, if also highly disturbing, reminder of a time long ago. But the use of science to define an ideal of purity in the human form lives on today, notably in the quest to identify and regulate the elite female athlete.

Not normal enough

In April 2018, the top international governing body for the sport of track and field—the International Association of Athletics Federations (IAAF)—released regulations aimed at limiting the participation of some female athletes competing at the international level in middle-distance running events. The Eligibility Regulations for Female Classification specifically target women with certain differences of sex development (DSDs) and with naturally occurring testosterone levels that exceed those of most other female athletes. To be eligible to compete, such female athletes must lower their testosterone with medication or surgery. This IAAF mandate, which requires unproven medical interventions in otherwise healthy individuals, has prompted considerable debate.

Biological sex is far more complicated than junior high school biology might suggest. Although most men have 46 XY chromosomes and most women have 46 XX chromosomes, biological science today recognizes that there are also 46 XX males and 46 XY females. The IAAF regulations apply only to female athletes with 46 XY sex chromosomes with certain DSDs and who compete in women’s running events of distances between 400 meters and one mile. The approach taken by the IAAF to

developing its latest version of female eligibility regulation is contorted and confusing. Earlier regulations released in 2011 focused on all women with high testosterone. These rules were suspended by the Court of Arbitration for Sport (CAS) in 2015, following a challenge by the Indian sprinter Dutee Chand, due to a lack of evidence on the relationship between naturally occurring testosterone and in-competition performance.

The next incarnation of the regulations was issued in April 2018 and focused on all women (that is, both 46 XX and 46 XY) with high testosterone resulting from DSDs, but only for the limited set of middle-distance events, justified by recently published IAAF research alleging that high testosterone was associated with elevated performance in these events. After one of us (Pielke Jr.), along with colleagues Ross Tucker of the University of Cape Town and Erik Boye of the University of Oslo, identified major errors in the data underpinning this research, the IAAF in February 2019 again changed the regulations, this time to focus only on 46 XY DSD females with high levels of testosterone, competing in the events from 400 meters to one mile. The IAAF explained that these, and only these, events are where 46 XY DSD individuals are known to compete, and this alone justifies the focus on these events.

Given confidentiality provisions, and the absence of systematic testing, it is unknown how many female athletes are affected by the regulations. The IAAF claims that over the past 10 to 15 years, perhaps 20 to 30 athletes at its biennial World Athletics Championships (out of about 8,000) have been 46 XY females with DSDs. Those very few women who have recently publicly acknowledged that they fall under the regulations are each women of color from nations of Africa, raising concerns about the role of race and nationality in the implementation of these rules.

One such woman is the South African 800-meter runner and two-time Olympic gold medalist Caster Semenya, who has been a target of IAAF regulatory efforts since she first became a World Champion as an 18-year-old in Berlin in 2009. Semenya was targeted because of her exceptional talent and, according to contemporaneous IAAF statements and those of some of her athlete peers, because of her appearance, which was deemed insufficiently feminine. In February 2019, Semenya appeared before the CAS in Lausanne, Switzerland, to appeal the latest IAAF regulations. The CAS ultimately upheld the regulations in a controversial ruling, in which the arbitral body acknowledged both the discriminatory nature of the regulations and a range of scientific opinion and concerns, but ultimately concluded that it was not within its mandate to revisit the IAAF's regulatory agenda, address human rights or medical ethics, or pass judgment on questions of scientific integrity. As a result, and pending a further appeal to the Swiss Federal Tribunal, Semenya and any other women who fall under the regulations are no longer eligible to compete

unless they comply with the requirement to lower their naturally occurring testosterone levels.

The IAAF regulations represent the most recent chapter in a much longer and controversial history of “sex testing” practices in sport, beginning in at least the 1960s. In track and field (called “athletics” outside the United States), the justifications historically offered by the IAAF for regulating female eligibility have emphasized “fairness” and preventing the inclusion of men in women's competition, since elite men perform better than elite women across Olympic events in track and field. Yet any effort to determine who is male and who is female is complex, since biological sex is not a binary attribute but occurs on a spectrum. As the historian Alice Dreger has written, “Humans like their sex categories neat, but nature doesn't care. Nature doesn't actually have a line between the sexes. If we want a line, we have to draw it on nature.”

The politics of reclassification

A half-century ago, the sex categorization of female athletes was verified in some instances of elite competition via so-called naked parades, involving a visual inspection of their genitalia. When this demeaning practice was abandoned, sport organizations adopted methods that they believed held the promise of scientifically and objectively telling us *what is*, rather than *what ought to be*, when defining the eligible female athlete. However, the promise of objective science has proven far more illusory than real, as the complexities of human biology have defeated all medical tests proposed by sports organizations to reliably divide biological sex into two distinct categories.

A closer look at the IAAF's current regulatory regime shows that the sports world hasn't come that far from naked parades, with science providing a fig leaf of respectability for practices that continue to demean and stigmatize certain women.

Before proceeding further, it is essential to dispense with one issue. The IAAF regulations discussed here are entirely separate from the rules that govern the participation of trans women in elite athletics (there are currently no regulations for trans men). These rules, implemented by the International Olympic Committee, define trans women as a separate category from DSD women since individuals in the latter category have experienced a continuity of gender assignment and identity from birth. Our focus here, like the IAAF regulations, is on 46XY DSD female athletes and whether a sport federation should have the authority to question and reclassify the sex of such athletes or require them to undergo medical treatment in order to compete.

The IAAF initially argued upon release of the 2018 regulations that it was *not* seeking to make a determination of gender or sex. Rather, it was merely regulating eligibility *within* the female category, by drawing a line between female athletes with “normal” sex development and other women with different development (for example, some 46 XY DSD

women lack ovaries and may have higher levels of testosterone than most other women). With this line of argument, the IAAF sought to distance itself from earlier, failed regimes of sex testing or gender verification, which had been severely critiqued in terms of ethics and science for seeking to reclassify the sex of some female athletes.

However, immediately before Semenya's appeal to the CAS in February 2019, the *Times* (London) reported that the IAAF had made a late change in its approach to the regulations and was indeed reclassifying some women as "biological males" based on their chromosomal make-up. The IAAF initially denied taking this approach in a press release responding to the media report. Yet when testifying during the CAS hearing and subsequently in public discussions, IAAF officials admitted that its regulations are based on the premise that some women are not in fact female but are instead biological males. In the words of the IAAF, the athletic ability of such athletes is elevated to such an extent relative to so-called normal women that their presence in certain female events is "category defeating." Further affirming the IAAF's goal of reclassification, the regulations state that such women are eligible to compete only in the male category unless they undergo medical treatment to reduce their naturally occurring testosterone levels to within the "normal" female range.

The biology of human sex development is fascinatingly complex. One group of female athletes targeted by the regulations, and discussed during Semenya's appeal to the CAS, are 46 XY women with a genetic variation known as 5-alpha reductase deficiency, type 2 (5-ARD2). Women with DSD conditions leading to elevated testosterone, but with XX chromosomes, are exempt from the regulations. Those 46 XY women with DSDs are also exempt as long as their testosterone does not exceed a certain threshold. Relative to other women, 46 XY 5-ARD2 females often have higher levels of testosterone. However, they also typically have insufficient levels of another hormone—dihydrotestosterone—to experience typical male development, hence their clinical classification as females. Thus, when the IAAF determines that some 46 XY females should be in fact be considered biological males, it misrepresents basic biological understandings and deviates from the widely shared position of the international medical community, such as reflected in statements by the World Health Organization, which recognizes XX males and XY females in addition to still more chromosomal variations.

When the CAS upheld the IAAF regulations in 2019, it was agnostic toward the IAAF argument that certain women athletes could be reclassified as biological males. The CAS found that it "does not consider it necessary specifically to

The use of science to define an ideal of purity in the human form lives on today, notably in the quest to identify and regulate the elite female athlete.

determine whether the IAAF's invocation of the concept of a 'male sport sex' possessed by 'biological males' and a 'female sport sex' possessed by 'biological females' is valid and/or proper." The avoidance of this point by the CAS is one of the perplexing elements of the ruling, given that the very basis of the IAAF regulations, expressed openly at the hearing by IAAF officials, is the argument that certain female middle-distance athletes are in fact not female at all.

Having tried to dodge the question of sex determination, the CAS panel returned to the prior IAAF rationale and endorsed conceptualizing the female category as divisible into those with "normal" sex development (and thus implicitly equated with "normal" athletic ability) and those with XY chromosomes and high testosterone levels. The CAS ultimately concluded that "female athletes with 5-ARD2 and other 46XY DSD have high levels of circulating testosterone in the *male range* and ... this does result in a significantly enhanced sports performance ability" over other women. Thus, testosterone levels were alleged by the CAS (and the IAAF) to be both sexually dimorphic and the overriding basis of female-male differences in middle-distance running ability, with both points being heavily debated during the Semenya appeal.

Trying to keep things simple

The IAAF regulations, and the CAS endorsement of them, are underpinned by the notion that women and men should be characterized by nonoverlapping distributions of testosterone. This argument was made by Stéphane Bermon, the director of the IAAF's Health and Science Department and a chief architect of the group's most recent regulations, during a June 2019 symposium at the French Embassy to the United States in Washington, DC. At the event, Bermon relied on an August 2018 literature review whose lead author, Richard V. Clark, is on the board of directors of the US Anti-Doping Agency. This study, published in *Clinical Endocrinology*, claimed to show "large divergence" in the testosterone ranges of "normal, healthy males and females." The authors present individuals with certain 46 XY DSD conditions as having testosterone levels that overlap with "normal, healthy males" and not "healthy females." Bermon explained to the symposium that the distribution shown in this study reveals that female 46 XY DSD athletes with high testosterone levels are in fact "biological males."

As with the case of Norma, the study by Clark and colleagues—and the IAAF in its use of it—purports to be presenting *what is*, rather than *what ought to be*. The paper states: "The purpose of this commentary was to summarize the *well-established reference range* of serum testosterone levels in normal, healthy adult males and females" (emphasis added).

Just as with Norma, however, what the Clark study establishes as “normal” and “healthy” is a function of choices made by researchers about who to include in calculating the summary statistics for such a population.

Let’s take a closer look. The Clark study reviewed 26 papers published in peer-reviewed journals in order to compile testosterone levels for “normal, healthy males and females” and several types of DSDs, including 46 XY 5-ARD2. Critically, the classification procedure used in the review begins by excluding people with DSDs from the definition of “normal” and “healthy” individuals.

After providing the testosterone ranges reported for 46 XY males and 46 XX females (eight studies), Clark and colleagues separately reported testosterone ranges for each of the following three groups of individuals with certain DSDs: “46 XY individuals with 5ARD2” (six studies), “46 XY individuals with PAIS/CAIS [partial and complete androgen insensitivity]” (seven studies), and “Females with PCOS [46 XX with polycystic ovary syndrome]” (four studies). Only females with PCOS were classified by sex at the outset. For both groups of 46 XY DSD individuals, the authors grouped males and females together. The researchers then reported testosterone ranges from the selected literature in the form of a forest plot (see Figure 1), which shows no overlap between those they classified from the outset as “normal and healthy” XX females and “normal and healthy” XY males. As shown below, the plot also displayed the reported testosterone ranges for the three DSD groups, with the authors placing two of them in the XY male column and one in the XX female column, based on chromosomes rather than sex reported in the reviewed studies, and showing for each a range of testosterone values that approximates healthy males and healthy females respectively.

The review concludes that there is a “marked, bimodal distribution of testosterone levels between males and females without any overlap” and that “there is no continuum of testosterone levels from normal females to normal males.” The authors further argue that “individuals with 46 XY DSD due to 5ARD2 are genetic males who as adults typically have serum testosterone levels within the normal adult male range.” Then, without providing any additional evidence, they lend support to the IAAF regulatory efforts by suggesting that “in adult *genetic males* with 5ARD2, elevated endogenous testosterone levels are likely associated with enhanced athletic performance relative to *genetic females*” (emphasis added). Thus, despite the fact that chromosomal tests—first used by sports organizations for sex testing in the 1970s—were abandoned because the genetic complexity of humans is not readily amenable to binary female-male categories, here they are again.

The methodological circularity of the review article should be obvious. First, the study separates out DSD individuals, whose inclusion in an initial classification would greatly

complicate the production of “clean,” nonoverlapping testosterone numbers for female and male categories. Instead, the authors present testosterone ranges for DSD individuals separately, suggesting they are other than normal and healthy, and unclassified by sex, despite the fact that each of these individuals is indeed already recognized as either female or male in the reviewed studies. The authors then use their bimodal “physiological reference framework” (or reclassification framework), developed from the preliminary groups of 46 XX females and 46 XY males, to reclassify DSD individuals as either female or male based on their chromosomes. This methodology is identical in form and application to the creation of Norma and then her use as an ideal to judge the broader population.

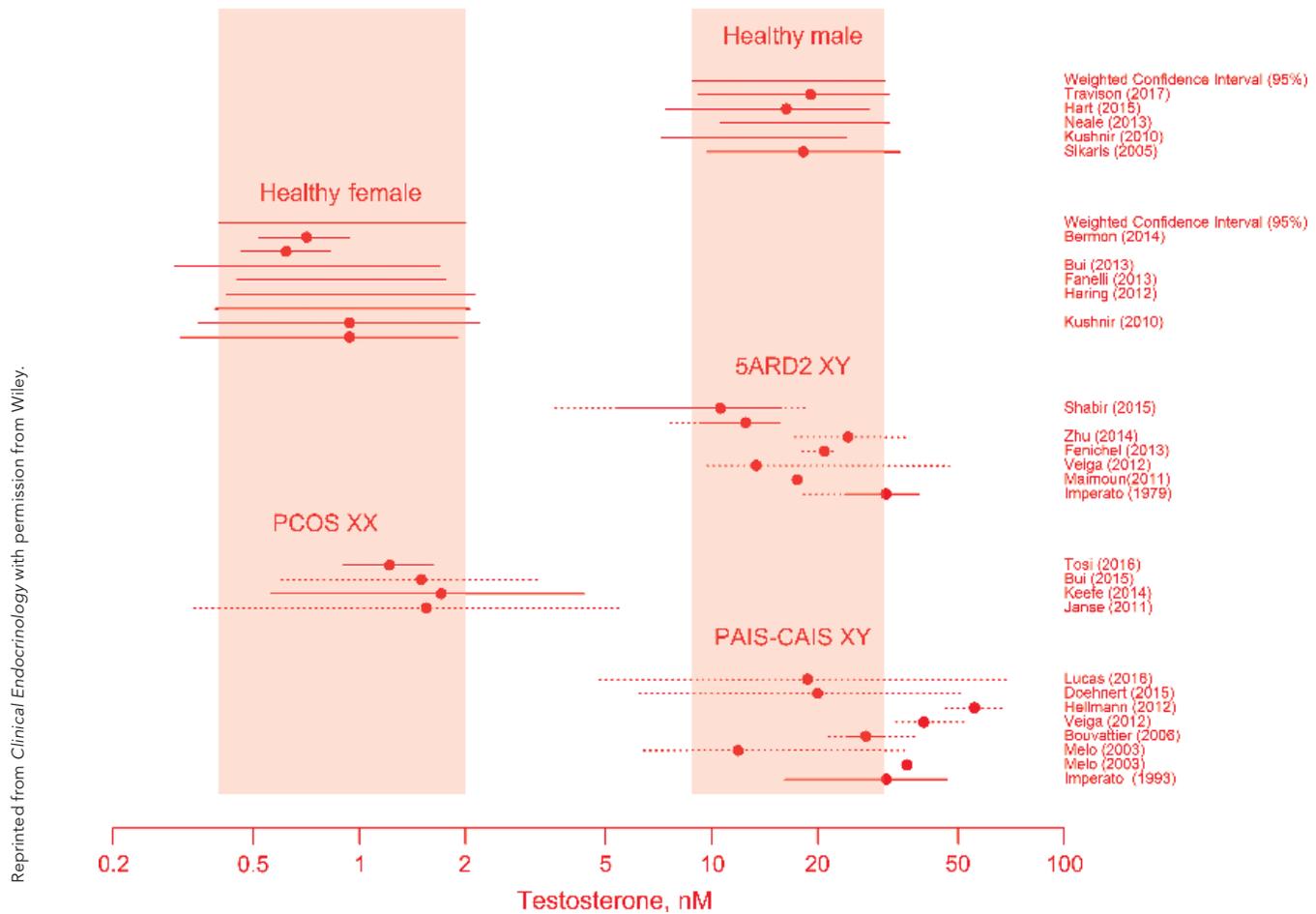
The circularity of this method is not unique to the 2018 study; it applies to any study that employs a pre-study sex classification of study subjects and then uses the resulting statistics to reclassify individuals who are outside the study population. The IAAF cites such studies in the regulations, invoked them before the CAS, and emphasizes them in its publications as the basis for using female and male testosterone levels for sex classification. The IAAF thus imposes the norms established by the researchers—the initial subjective judgments of what membership in a given category should look like—onto the data, telling us not *what is*, but *what (according to the investigators) ought to be*.

Whether 46 XY DSD individuals are either female or male depends not on testosterone levels, or even on chromosomal make-up, but on the sex assigned to them at birth, based primarily on an examination of their genitalia and maintained from that moment forward (or not) depending on how their gendered lives unfolded. For example, several of the studies included in the review by Clark and colleagues, which assessed testosterone ranges for the 46 XY DSD 5-ARD2 category, identified these individuals as either female or male. The methodology used in the Clark study ignores this fact and instead defines them collectively and principally as unhealthy, abnormal, and with a questionable sex classification.

A rather bizarre consequence of this approach is that 46 XY DSD individuals who are perfectly healthy, including female athletes competing at the elite level of international track and field, are deemed unhealthy. The methodology also conceals the reality that considerable testosterone variation across individuals classified as female or male from birth can be considered a biologically, if not statistically, normal occurrence, even if the DSD conditions are relatively rare.

The problems with the Clark study are, however, more than just methodological: there are substantive problems as well. In the process of replicating the study’s literature review, we found several major errors in the reporting of data from the reviewed studies, most notably the failure to report the full range of data from one of the reviewed studies of 5ARD-2

Fig 1. Original, Erroneous Forest Plot from Clark et al.



Testosterone ranges for “normal, healthy” females and males and three different DSD groups.

individuals. After we notified the authors and journal of these errors, *Clinical Endocrinology* published a lengthy erratum that included a revised forest plot with the corrected values (see Figure 2). The corrections now reveal an overlap between the testosterone ranges of 46 XY 5-ARD2 individuals and both the “healthy male” and “healthy female” categories. Contrary to the conclusions initially reported and highlighted by the IAAF, the use of testosterone combined with chromosomal attributes in an effort to create distinct male and female categories is not only a reflection of subjective methodological choices but also fails to support the original conclusions.

Inclusivity instead of circularity

What might an alternative methodological approach to classification look like?

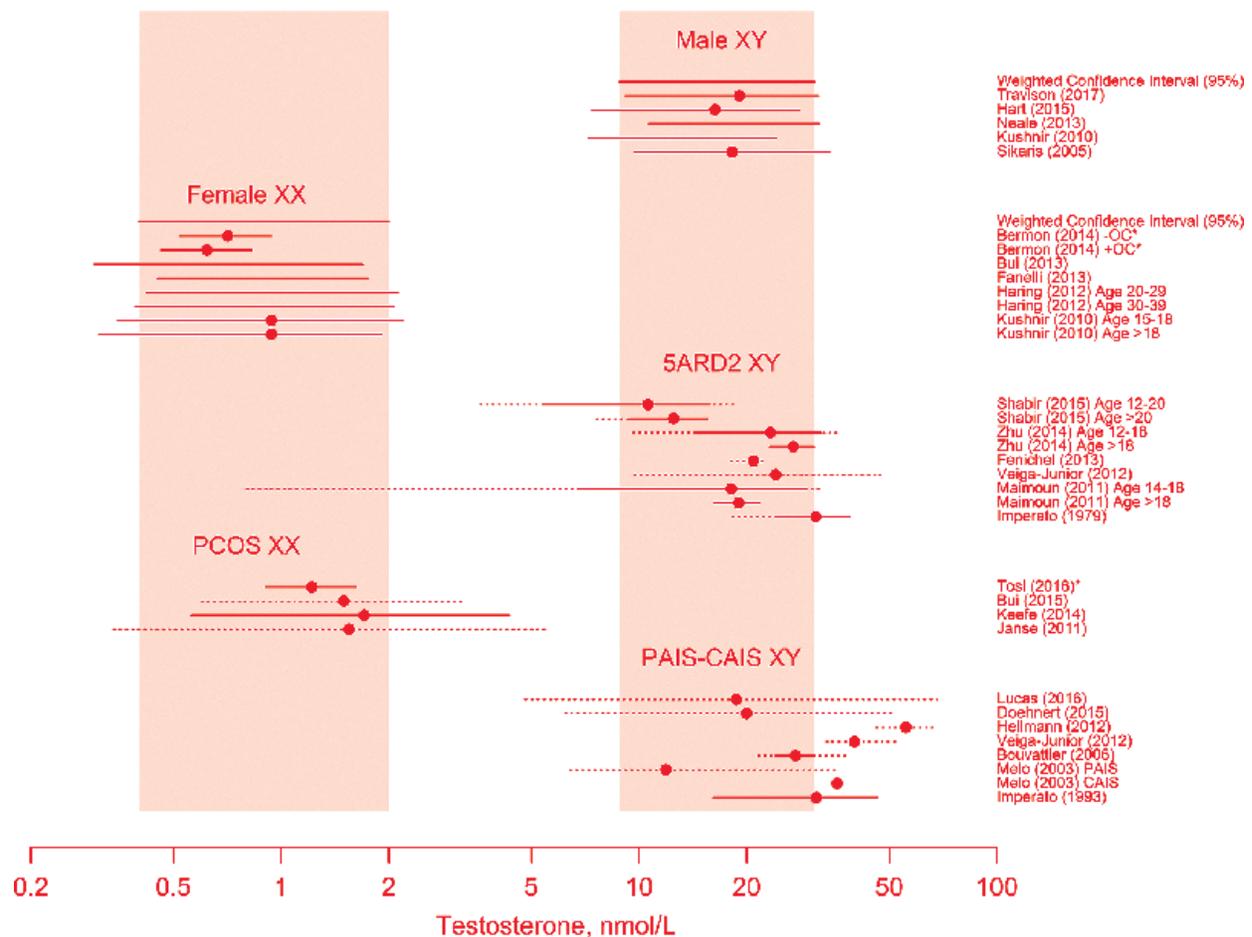
Instead of relying on the variability-reducing constructs of “normal” and “healthy” to justify excluding certain individuals

(including world-class athletes) from the initial study pool, an alternative approach to classification would:

- First, include all individuals in the initial study population, whether with DSDs or not;
- Second, classify them as “male” and “female” based upon their sex assigned and maintained from birth;
- Third, only then assess testosterone levels within each group; and
- Fourth, present these female and male ranges without a separate classification of DSD individuals.

Although full implementation of this methodology is beyond our scope here, in Figure 3 we show testosterone ranges from two of the studies reviewed in the Clark study according to the sex of the individuals as reported by these papers. Of individuals who are 46 XY 5-ARD2, according to

Fig 2. Revised Forest Plot from Clark et al.



Reprinted from Clinical Endocrinology with permission from Wiley.

Correct testosterone ranges for 46XY DSD 5ARD2 individuals.

one of these studies, approximately 30% identify as female. The figure clearly shows that if one classifies the testosterone ranges of 46 XY 5-ARD2 individuals based on their actual sex, there is a complete overlap between males and females, as well as overlap with the testosterone ranges of the “normal” and “healthy” females that Clarke and colleagues used to develop their reclassification framework.

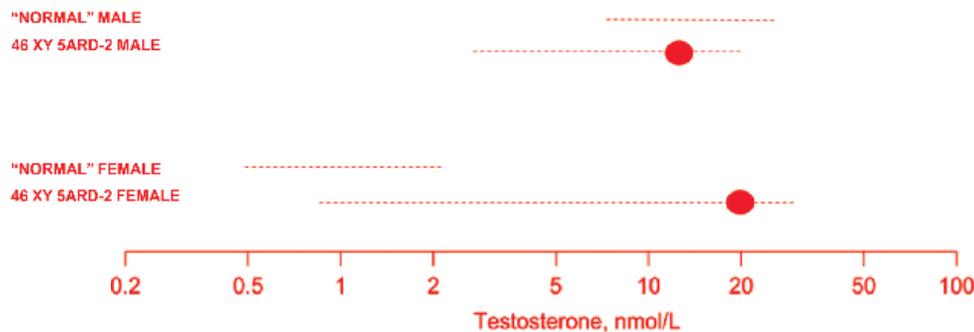
The choice to be inclusive of DSD individuals in study design (as we recommend) or exclusive of these individuals (as in the Clark study) is fundamental to the results. Here, as with Norma, it is the pre-study decision-making that determines who is deemed ideal and who is not. Ultimately, when such decisions are portrayed as scientific rather than subjective, they can reinforce discrimination by making categories seem like entirely natural phenomena rather than a mix of the natural and the social.

In the end, either approach—to exclude or include certain individuals from the initial classification—is a subjective

choice. Science does not determine this choice. Both approaches could be claimed to be scientific and evidence-based. But the point to emphasize is that science and data are not doing the work here: choice of methodology leads to diametrically opposed results. Under the methods used in the 2018 study, which appears to have been a foundation of the CAS decision, Caster Semenya, a female since birth, would be reclassified as a male. Indeed, in the lengthy correction to the Clark study, after the revised testosterone ranges offered less support to their claims of a clear demarcation, the authors introduced a new methodological step not found in the original paper: they simply defined all 46 XY individuals as male, regardless of whether they were reported as female in the reviewed studies. By defining 46 XY 5-ARD2 individuals as male, the authors simply assert what they had initially set out to prove with evidence.

Under our alternative classification methodology, Caster Semenya would be classified as a female, as she has been since

Fig 3. Testosterone Ranges of 46 XY 5ARD-2 Individuals



Testosterone ranges of 46 XY 5ARD-2 individuals as reported in two studies reviewed by Clark et al., according to the sex of individual as reported in the studies, along with testosterone ranges (the 2.5%-97.5% intervals) of "normal" males and "normal" females as reported in Clark et al. Median values are shown for the 46 XY 5ARD-2 males and females, but are not reported for "normal" males and females by Clark et al.

birth. Similarly, the subjects of the various studies reviewed in the 2018 Clark study would be classified based on their sex assigned and maintained from birth. Statistics do not provide an objective answer to how classification methods are to be employed, but they can be wielded to give the impression that they do. Science alone is unable to determine the boundaries of the female category, either on or off the track.

Efforts to render variation invisible or abnormal and to reclassify those female athletes who disrupt the IAAF's preferred construction of sex presume mutually exclusive female-male categories with biological traits that are distinct (XX female versus XY male chromosomes) and nonoverlapping divergent testosterone ranges. Importantly, this binary world is not what is, but what the IAAF believes ought to be.

Modern track and field (and many other sports) is organized around binary definitions of male and female that evolving science and gender politics have rendered more complex, fuzzy, and ambiguous. The Caster Semenya story is thus yet another example of the difficulties that social institutions have in adjusting to shifts in both gender politics and scientific knowledge. But it is also a story of misguided institutional expectations: the IAAF has sought to deliver clarity and certainty by invoking "science" as the basis of its decision-making, but when it comes to biological sex, science in fact delivers the opposite.

Such a realistic view of science should be viewed as an opportunity. Rather than being "category defeating," as the IAAF has argued, the alternative classification methodology that we propose is in fact "category reinforcing." Our approach maintains female-male competition categories. It allows these categories to be retained in a form that reflects the actual biological complexity of sex and the heterogeneity among female athletes while also respecting their biological sex as assigned and maintained since birth. Our approach has the advantage of not empowering sports organizations to reassess and potentially reassign female classifications, much less mandate a requirement for unproven and unethical medical interventions. The Women's Sports Foundation and the

International Working Group on Women and Sport agree: they have argued that women have much to gain from a more inclusive approach, since existing regulations discourage excellence among female athletes based on naturally occurring traits and encourage the scrutiny and regulation of female bodies. For the IAAF, Caster Semenya and other women with genetic variations are abnormal and must be excluded unless they medicate to remedy their imperfections. Our view is that Caster Semenya is already perfect, just as she is.

Roger Pielke Jr. (pielke@colorado.edu) is a professor at the University of Colorado, Boulder. **Madeleine Pape** is an Olympian and a postdoctoral fellow at Northwestern University.

Recommended reading

- Richard V. Clark, Jeffrey A. Wald, Ronald S. Swerdloff, Christina Wang, Frederick C. W. Wu, Larry D. Bowers, and Alvin M. Matsumoto, "Large divergence in testosterone concentrations between men and women: Frame of reference for elite athletes in sex-specific competition in sports, a narrative review," *Clinical Endocrinology* 90, no. 1 (2019): 15–22.
- Doriane Lambelet Coleman, "Sex in Sport," *Law and Contemporary Problems* 80 (2017): 63–126.
- Kathryn Henne and Madeleine Pape, "Dilemmas of Gender and Sports Governance: An Invitation to Southern Theory," *Sociology of Sport Journal* 35, no. 3 (2017): 216–225.
- Rachel Mulot, "Athletism: 'If you want to compete in the feminine category, then you must not oppose a treatment,'" *Sciences et Avenir* (June 25, 2019).
- Madeleine Pape, "Expertise and Nonbinary Bodies: Sex, Gender, and the Case of Dutee Chand," *Body & Society* (2019).
- Roger Pielke Jr., "Sugar, spice and everything nice: how to end 'sex testing' in international athletics," *International Journal of Sport Policy and Politics* 9, no. 4 (2017): 649–665.
- Roger Pielke Jr., Ross Tucker, and Erik Boye, "Scientific integrity and the IAAF testosterone regulations," *The International Sports Law Journal* (2019): 1–9.