Published in final edited form as:

Front Neuroendocrinol. 2020 January ; 56: 100817. doi:10.1016/j.yfrne.2019.100817.

The Promises and Pitfalls of Sex Difference Research

Liisa AM Galea¹, Elena Choleris², Arianne YK Albert³, Margaret M McCarthy⁴, Farida Sohrabji⁵

¹Department of Psychology, Djavad Mowafaghian Centre for Brain Health, University of British Columbia, Vancouver, BC V6T1Z3, Canada

²Department of Psychology and Neuroscience Program, University of Guelph, Guelph, ON N1G 2W1, Canada

³Women's Health Research Institute of British Columbia, Vancouver, BC, Canada

⁴James and Carolyn Frenkil Dean's Professor and Chair, Dept of Pharmacology, Univ of Maryland School of Medicine, Baltimore, MD, United States

⁵Women's Health in Neuroscience Program, Department of Neuroscience and Experimental Therapeutics, Texas A&M HSC College of Medicine, Bryan, TX 77807, United States

Abstract

Funding agencies in North America and Europe are recognizing the importance of the integration of sex differences into basic and clinical research. Although these mandates are in place to improve our knowledge of health for both men and women, there have been a number of implementation issues that require vigilance on the part of funders and the research community. Here we discuss issues on simple inclusion of both sexes in studies to specialisation of sex differences with attention paid to statistics and the need for sex-specific treatments. We suggest differing mandates need to be considered regarding simple integration versus the need for studies in the specialisation of sex differences and/or the need for research that recognises the importance of male-specific or female-specific factors that influence subsequent health such as menstruation, menopause or pregnancy.

Keywords

sex differences; implementation; sample size; statistics; females; males; funding; SABV; SGBA

Diversity in research matters: It is vital to study both sexes but is it vital to study sex differences?

In 2016 the USA National institute of Health (NIH) mandated that all preclinical research must include sex as a biological variable (SABV), unless strongly justified otherwise. This

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

has substantially increased attention to sex differences research, with good reason. Studying how biological sex contributes to our health can help understanding of disease etiology, manifestation, progression, and treatment of disease. Indeed, males are more likely to be diagnosed with autism spectrum disorders or develop Parkinson's disease whereas females are more likely to be diagnosed with major depressive disorder, anxiety disorders, autoimmune disease, and multiple sclerosis (Brookmeyer et al., 1998; Golden and Voskuhl, 2017; Gutiérrez-Lobos et al., 2002; Irvine et al., 2012; Kessler and Bromet, 2013; McPherson et al., 1999; Sohrabji et al., 2016). Sex differences are also noted in specific disease subsets, thus more females than males display the relapsing-remitting type of multiple sclerosis than men (Golden and Voskuhl, 2017). Similarly, whereas the incidence of epilepsy is higher in males than females, only women are susceptible to a catamenial epilepsy, where seizures are tied to the menstrual cycle, and up to 70% of females with epilepsy show a variant of catamenial type (Reddy, 2017). Perhaps less well known, there are also sex differences in the timing or onset of neuropsychiatric disorders. Onset of obsessive compulsive disorder is more likely to occur in early adolescence for males but during the perinatal period for females (Mattina and Steiner, 2016). Even when prevalence of disease shows no sex bias, such as in schizophrenia, there can be profound differences in timing onset. Males are more likely to present with schizophrenia as teens, but females are more likely to present a couple of years later than males with a secondary peak in middle-age (Häfner et al., 1992). Each of these examples gives us important clues on the nature of the disease. Yet, sex differences in onset and prevalence are rarely explored despite that studying these sex differences could yield powerful clinical and preclinical models of disease, and clues to disease etiology and pathology.

Sex differences are also found in the presentation of a variety of diseases (De Bellis et al., 2019; Golden and Voskuhl, 2017) and in mechanisms surrounding disease such as in neuropathic pain (Sorge et al., 2015). Schizophrenia is associated with lower temporal lobe volume (Bryant et al., 1999) and reduced white matter fractional anisotropy in the left cingulate in males but not in females (Lang et al., 2018). Studying differences in manifestation and mechanisms of disease both within and between the sexes is critical to strive for efficacious treatments. Indeed, the lack of consideration of sex has led to a stalling in new treatment discoveries (described below for stroke) and fields that have incorporated sex differences into medical practice have led to better health outcomes for both males and females. A good example of this is the decrease in mortality due to cardiac disease in both males and females across the globe (Menash et al 2017). In the US, this success can be attributed in part to the landmark Framingham Heart study which began enrollment in 1948. The original cohort of 5000+ participants was 55% female, which was in stark contrast to contemporaneous studies where females were either underrepresented or excluded altogether (reviewed in Mahmood, et al, 2014)

Great strides in discovery have been made in sex-specific diseases; survival rates have increased 16% to 33% over the last 35 years in female-only breast cancer or in prostate cancer (<https://seer.cancer.gov/statfacts/>). The five-year survival rate is an astounding 90% for female-only breast cancer and 98% for prostate cancer. Contrast these with survival rates in cancers that affect both sexes. Prevalence of lung and bronchus cancer is 19.4% but unfortunately the increase in survival rate is only 8% over the last 35 years. There are many

reasons for these larger improvements and better survival rates in single-sex cancers, but it is hard not to see that real progress can be made when one sex is studied for a sex-specific disease. The same argument holds for studying disease with a lens on sex as a biological variable. However when there has been little success in drug development for certain sex-biased diseases such as major depressive disorder and Alzheimer's Disease, it is plausible this may be due in part to the lack of consideration of sex. Advances will be made if researchers are mindful that sex-specific treatments are sometimes necessary. A comprehensive analysis of data from patients with glioblastomas, a form of malignant brain tumors (Yang et al., 2019), found different therapies were more successful in females than in males, and that this matched up with molecular markers in the tumours themselves. In the case of experimental stroke, there are now several examples of drugs that are effective in one sex and not the other (Sohrabji et al., 2017), compelling us to acknowledging that sex may be a critical variable in drug efficacy. This is well aligned with the goal of "precision medicine". Females are more likely to suffer from side effects from pharmacological agents (Haack et al., 2009) but also from surgery (Giustino et al., 2019; Previato et al., 2018), and therefore more likely to be under-prescribed pain medications or other therapeutics. Health disparities between the sexes are what arguably led the NIH to mandate the inclusion of both sexes in clinical trials that was extended to preclinical research in 2016.

We now must ask, will mandating SABV effectively change our understanding of health and disease? Unfortunately the NIH mandate on clinical trials did not extend to presentation of the data by sex nor did it mandate sample size and thus there has been little progress to date in terms of reporting and analysing outcomes by sex. Furthermore, NIH-sponsored trials account for only a small (~5%) subset of all clinical trials (Ehrhardt et al., 2015). As laudable as these mandates were, they have done little to move the dial, due, in part, to lack of integration of analyses of sex or reporting of the data by sex in these trials (Geller et al., 2018). Thus, even though researchers were required to include both sexes in these trials, the majority have not analysed the data with sex as a consideration (Geller et al., 2018; Labots et al., 2018). A recent study asserted there is no systematic under representation of females in clinical trials registered with the Food and Drug Administration (Labots et al., 2018). This was certainly true of the 38 of 137 drug trials they examined, but 72% of these drug trials did not analyse or report the outcome data by sex (Labots et al., 2018). Another study concluded that only 26% of 107 NIH-sponsored studies in 2015 included sex as a variable in their statistical analysis (Geller et al., 2018). However, this included studies that used sex as a covariate. The use of a covariate is statistically removing the influence of sex as a (linear) factor, thereby effectively removing the reason for the mandate. Mersha et al. (2015) studied the effects of sex-combined analyses on revealing Single Nucleotide Polymorphisms (SNPs) related to asthma. There are several SNPs and differentially expressed genes from different tissues (Gautam et al., 2019), that show large sex disparities. In the case of asthma, 47 SNPs were detected using a sex-stratified analysis, however when both sexes were combined only 21 SNPs were identified. Thus, over 50% of the SNPs were lost when sex as a variable was not included in the analysis. Until studies are mandated to report their data disaggregated by sex and analysed for the effect of sex, and/or make the data freely available to all, the inequities will continue. Although the authors acknowledge that incorporating sex differences is important, studying sex differences requires nuanced approaches and

considerations to ensure successful transition of the SABV mandate into fruitful discovery and are discussed below.

Statistics, statistics, statistics

The funding agencies, such as NIH and the Canadian Institutes of Health Research (CIHR), which are mandating inclusion of both sexes are, for the most part, not providing additional funds to study both sexes. As such researchers are naturally resistant to the idea that they need to double their sample size and perform the same research with effectively less money. The funding agencies have responded by indicating that sample size does not need to be doubled or even changed to investigate sex differences in research (<http://www.cihrirsc.gc.ca/e/51257.html>). However, this assumption is problematic on a few levels. A recent paper showed that doubling of sample size was not necessary if one did not predict main effects of sex (Busch et al., 2019). However, if one has never tested a drug's effects in both sexes it would be impossible to predict whether main or interaction effects may be expected. For best practices, it is important to ensure that studies are statistically powered for interactions with sex differences to be detectable. There are numerous instances in neuroscience, and other areas, of underpowered studies (Button et al., 2013), and it does not make statistical sense to assume the same sample size will be informative (e.g. the advice is if you previously used a sample size of 6 males, you could instead use 3 males and 3 females). Perhaps a doubling of sample size is not required but more subjects will be required to have enough power to inform on a sex difference if one exists.

The need for larger sample sizes is because in general interaction effects require larger numbers to detect than main effects (Cohen, 1988). The advice above to use the same number of animals split between the sexes is predicated on two implicit assumptions 1) that the numbers will be split equally between males and females (balanced design), and 2) that any sex difference in the effect of the main intervention under examination is either extremely large or non-existent. Assumption one should be feasible to achieve in practice. However, the second assumption for the most part would prevent the detection of smaller but still important differences in effects between the sexes. We conducted some simple simulations of experiments using 2-way ANOVAs with sex and an intervention as the between-subjects variables. This would result in examination of two main (overall) effects on sex and the intervention and an interaction effect between the intervention and sex. An interaction effect would be significant if the intervention had a different effect size in males vs. females. By varying the size of the interaction effect we determined the sample size actually required to detect it as significant with 80% power and $\alpha = 0.05$, assuming equal allocation between both treatment arms and sexes (fully factorial design). We investigated 4 possible scenarios: 1) effect of treatment is exactly reversed in males and females (very strong interaction effect), 2) the treatment has no effect in females but does in males (a large interaction effect), 3) the treatment is half as effective in females compared to males (a moderate to large interaction effect), and 4) the treatment is $\frac{3}{4}$ as effective in females compared to males (a moderate to small interaction effect). While scenarios 1 and 2 had sufficient power (80%) with a sample size of 20 (10 males, 10 females), and in fact a smaller sample size could be used in scenario 1, we did not achieve 80% power for scenario 3 until a sample size of 100, and it took a sample size of 350 before achieving 80% power

for scenario 4 (Figure 1). It could be argued that a $\frac{1}{4}$ reduction of effect in females may not be relevant at the initial stages of examining a treatment effect, however, an effect that is half as large in females as males is relevant for future investigations and clinical applications. It is also true that the exact numbers estimated here are the result of the specific parameter values chosen (means and standard deviations (SD)) in our simulations. We did try to choose fairly large effects to reflect those often found in animal models (SD = 2 for all groups, mean in male controls = 10, mean in female controls = 8 (effect size of sex $d = 1$), mean in males with treatment = 15, means of females with treatment = 1) 3, 2) 8, 3) 10.5, and 4) 11.75). The sample size needed for a specific study would need to be estimated given the information available for that particular hypothesis.

Insufficient statistical power and use of combined sexes in studies is often used as a reason to not examine sex as a factor. However, the use of sex as a stratified factor in the analyses can lead to *increased* statistical ‘power’, particularly when there are interaction effects. In the example of SNPS associated with asthma, the effect sizes were increased when sex was included as a variable (Mersha et al., 2015), this will be true whenever the sexes show opposing effects as demonstrated above with our simulations. Opposing or interaction effects between the sexes are common in biomedical research. Temporal lobe volume decreases with schizophrenia in males but not in females (Bryant et al., 2014); intranasal oxytocin increases brain activity in social areas in males but decreases it in females (Rilling et al., 2012); acute stress increases CA1 spine density in males but decreases it in females (Shors et al., 2001). Thus, it is important that researchers not just consider that sex differences will result in overall (main) effects but that they may result in interaction effects (when a treatment has different effects in one sex versus another), a point that is lost in the recommendation to not increase overall sample size when including both sexes (Buch et al., 2019).

There are instances in the literature where failure to examine sex differences past main effects of sex may lead to abandoning promising outcomes for one sex or the other. Progesterone as a treatment for traumatic brain injury had strong evidence to support its use in clinical trials, yet the phase 3 clinical trial failed perhaps due to the inclusion of both sexes. Wright et al. (2014) showed a statistical trend for an overall (main) effect of sex ($p = 0.07$) yet the data indicate that placebo was significantly better in females than in males. This suggests that the use of progesterone as a treatment for TBI may be fruitful but only for males and perhaps only in certain cases when age or TBI severity is considered. Unfortunately, the authors do not show us these interactions, but exploring sex as a significant factor in an interaction may have provided evidence that progesterone is a viable treatment for TBI in subpopulations of patients.

All in all, the recommendation to not double sample size will have no value in interpreting studies and will actively hamper understanding of disease processes. It would be akin to saying that if there was an ‘age’ mandate, that it would be appropriate to have 3 young and 3 aged animals in the same group. Thus, we recommend that at least in the first series of studies there should be sufficient power to detect sex differences beyond main effects as suggested above. Further it is important to recognize that there are many types of sex differences, and that sex differences in the underlying neural mechanisms can be found even

when there are no apparent sex differences in the trait studied (Becker and Koob, 2016; Shansky and Woolley, 2016).

The Trouble with Hormones: Hormonal Cycles in Males and Females

Rodents typically have a 4–5 day estrous cycle, and primates, including humans have a 30-day menstrual cycle. In a ground-breaking study, Prendergast and colleagues (2014) found no differences in variation between the sexes on several physiological responses. These authors and others (Prendergast et al., 2014; Becker et al., 2016) have used these data to suggest that it is not necessary to examine estrous cycle. However, the lack of sex difference in within group variability does not imply that the variation within one or both sexes is not due to hormones. Indeed, in the original article, they point to differences in housing as the source of the variability in males. Male mice form dominance hierarchies, which are associated with testosterone levels (Williamson et al., 2017). Thus, the variation in male mice is likely to be due to dominance hierarchies that has at the very heart of it – hormones.

The idea that females are too complicated to study because of their hormones has been expertly argued to be sexist (Shansky, 2019). It is also equally valid to assert that because human males have diurnal fluctuations in testosterone levels (Harden et al., 2016), they too are complicated to study. Indeed, both males and females have diurnal fluctuations in glucocorticoids with high levels of cortisol released upon awakening, and lower levels throughout the day (Harden et al., 2016). These fluctuations are important for several circadian functions and have been tied to disease as well as symptom severity (Dominoni et al., 2016; Fonken et al., 2019; Tackett et al., 2015). Thus, clearly hormonal fluctuations are not simply a female phenomenon.

Aging in both sexes is accompanied by a reduction in hormones, which starts at an earlier age in human males. In humans, females show a large drop in 17 β -estradiol and progesterone serum levels at menopause, whereas male levels of testosterone decline gradually starting at 30 years of age (Harman et al., 2001). Furthermore, the ovaries do not cease to function at menopause as androgens continue to be released from the ovaries long after menopause in humans (Davey, 2012). Thus, the idea that hormones are only relevant to females is inaccurate, hormones modulate behavior in both males and females and to similar degree. Perhaps more interesting is not that hormones have diurnal or monthly variations, or that they may be tied to disease but that these variations are differentially affected by other hormone systems in a sex-specific manner. Intriguingly, the hypothalamic pituitary adrenal axis interacts with the hypothalamic pituitary gonadal axis in opposing ways in males and females as corticosterone levels can suppress testosterone levels in males but increase 17 β -estradiol levels in females, at least acutely (Goel et al., 2014; Viau, 2002). Diurnal variations in the occurrence of heart attacks are well known, with more heart attacks likely to occur between 6 am and noon as compared to 6 pm to midnight. Moreover, ‘morning’ heart attacks are more severe as judged by release of 2 enzymatic markers, creatine kinase and troponin-I (Suárez-Barrientos, A., et al., 2011). This was also seen in experimental studies of mice subject to myocardial infarction at different times of the day (Durgan et al., 2010), although neither study indicates if both sexes were used or if data was disaggregated by sex. There are

numerous other sex specific interactions with metabolism, immune and HPG systems that are underappreciated.

Pink Pills and Blue Pills

The requirement of funding agencies to include females in research comes from health advocates and researchers alike sounding the alarm that females have been less studied and as such, less is known about women's health. Women live longer, but suffer from more chronic illnesses, misdiagnoses, delays in diagnoses, and suffer from more side effects of drugs (Alberich et al., 2019; Leveille et al., 2000; Wada et al., 2019; Westergaard et al., 2019). Because much of the previous work has been done in males and treatments have been designed in male models, we have come a long way in understanding disease mechanism in males. Therefore, it is no wonder that treatments fail to work well in females, given basic research is dominated by work in males (Berry and Zucker, 2011; Wills et al., 2017). There is good evidence that when researchers take the mandate to incorporate SABV seriously, new important knowledge is generated (Sorge et al., 2015; Zeng et al., 2018). Studies found that the mechanisms for pain and immune control are different between the sexes (Sorge et al., 2015; Zeng et al., 2018). Bacteria capture is led by different mechanisms in males (complement opsonisation) compared to females (estrogen-driven antibodies) in the liver (Zeng et al., 2018), suggesting the need for different treatments in males versus females. This should not be so surprising given that the manifestation of disease can be quite different between the sexes (Wada et al., 2019; Eid et al., 2019; Young et al., 2018).

It should be clear by now that not considering both sexes in research is problematic for research in females but it is also problematic in males and two examples are given to illustrate this point. Zolpidem (trade name Ambien) was the first drug initially approved by the FDA with an equal dose for men and women. However, 20 years later the FDA recommended different doses for females compared to males (<https://www.fda.gov/drugs/drug-safety-and-availability/questions-and-answers-risk-next-morning-impairment-after-use-insomnia-drugs-fda-requires-lower>). Zolpidem administration resulted in greater plasma levels for longer in females compared to males (Greenblatt et al., 2000), thus impairing next morning mental alertness, even after correcting for body mass (Greenblatt et al., 2013). An instructive example of how inattention to sex differences can be catastrophic to men is illustrated by the research and eventual clinical trials of the compound called tirilazad mesylate. Tirilazad mesylate is a membrane lipid peroxidation inhibitor and free radical scavenger, which displayed remarkable anti-oxidant and neuroprotective properties for traumatic brain injury and subarachnoid hemorrhage (SAH). Tirilazad mesylate belongs to a family of 21-aminosteroid compounds whose effects in preclinical studies were so dramatic, they were dubbed 'Lazaroids' (reviewed in Cahill and Hall, 2017). This compound was subsequently tested in several clinical trials but was eventually not approved by the FDA. The first trial found that patients receiving tirilazad had reduced mortality and showed better scores on the Glasgow Outcome Scale as compared to the placebo group (Kassell et al., 1996), however the beneficial effects of the drug were restricted to male patients. This was followed by several more trials using this drug for patients with SAH or traumatic brain injury, all of which failed to show significant effects. It should be noted, that despite the intriguing sex difference seen in Kassell et al 1996 study, subsequent studies did not

incorporate SABV, although male and female patients were included in all but one case. Thus, the drug was never approved, and a promising drug was permanently tabled. In subsequent studies, several labs have shown in preclinical studies that drugs that are effective for stroke in one sex may not be effective in both sexes. These are missed opportunities that are potentially costing lives of men and women worldwide.

Women's Health is not just how women are different from men

Female-specific experiences such as pregnancy/postpartum, menstruation, menopause, and hormonal contraceptive use (Gierisch et al., 2013; Skovlund et al., 2016) influence health. Parity influences cardiovascular disease, metabolic disease (Kim and Lee, 2017), and Alzheimer's disease (Jang et al., 2018; for review see: Galea et al., 2018). However, studies are scarce in how female-specific experiences can alter the trajectory of female health. These experiences can influence disease prevalence but also possibly treatment. However, studying sex differences exclusively will not address how these experiences can impact health and it begs the question does NIH's SABV and CiHR's sex and gender based analysis (SGBA) go far enough. Do we always need to study males in addition to females? Female health is not just about how it differs from male health. Oral contraceptives increase antidepressant use in adolescents (Skolund et al., 2016) but may be protective for depressive symptoms in adult (25–34) females (Keyes et al., 2013). Parity increases the risk for cardiovascular disease, osteoporosis, metabolic disease but decreases the risk for a variety of cancers (Hunkula et al., 2006). Studying these effects can lead to new treatments for females but they may also give us clues for new pathways to investigate in males. However, studying fatherhood in males is not the same or even equivalent to studying it in females. Particularly in animal models, most mammalian species are not biparental and thus, adding males into the research is not scientifically meaningful. As male mice and rats can be infanticidal, males are not housed with pregnant and nursing dams (Chen et al, 2019), thus their experiences are very different than females. It is not always warranted to study both sexes, and not necessary to compare females against males. In addition, it may be that the swing towards SABV and SGBA may in fact be negatively impacting female only studies. Will et al (2017) examined studies in mice and rats from 2010 to 2014 in 6 journals. They found that overall, female only studies in these 6 journals accounted for only 6% of studies in 2014 whereas male only studies accounted for almost half of the studies at 40%. The good news was that the percentage of papers that did not specify sex decreased by 28% and the number of papers using both sexes increased by 14% over this 4-year period. An unfortunate consequence of the increased reporting was that the number of male-only studies increased by 9% while female-only studies only increased by 1%. To benefit both men and women's health and provide best treatments, sex needs to be considered but so do sex-specific factors. For women's health menopause, oral contraceptive use, gestational syndromes and parity matter for disease risk and treatment. Funders, policy makers, and reviewers should recognize that single sex studies are also important, and funding for women's only (and men's only) health specific questions should reflect this issue.

Implementation Issues: If Funders offer RFAs, researchers will come

Implementation issues are starting to arise from SABV and SGBA mandates. There is large variability in the ability of researchers to integrate sex as a biological variable and in reviewers in understanding proper integration of SABV in their reviews (Woitowich and Woodruff, 2019). Obviously, it takes training to become an expert in the study of sex differences beyond simple incorporation of both sexes in research and expecting reviewers and/or researchers to be well versed in the practice will be difficult even with excellent training tools. CIHR is doing its best to train reviewers and researchers alike to understand the importance of SGBA (Tannebaum and van Hoof, 2018). In an analysis of reviewers at NIH from 2016–17 Woitowich and Woodruff (2019) found that 68% of the reviewers who responded to the survey thought that SABV was an important consideration in 2017. But note that 32% of NIH reviewers did not think SABV was an important consideration and that there was only a 5% decrease in those percentages in the past year. Given the low rate of response (15% of all NIH reviewers) and likely selection bias (perhaps reviewers that thought SABV was important were more likely to respond), this may underrepresent reviewers that are not convinced as to the importance of SABV. Only 42% of NIH reviewers thought SABV would improve rigor and reproducibility and, disturbingly, there were large discrepancies in whether reviewers used SABV as a scoreable issue. The same researchers suggest that grant applicants may still not be adding females to their studies (Woitowich and Woodruff, 2019b). Perhaps more concerning, 88% of these reviewers felt confident in their understanding of the SABV policy, which begs the question of why there were questions on how to interpret SABV as scoreable issue and that almost half did not think SABV would improve rigor and reproducibility. These findings suggest that the funders have work to do to convince researchers that sex matters in health research (Woitowich and Woodruff, 2019c); and need to convince reviewers and researchers alike that sex differences need to be studied properly with attention paid to sample size and analysis plans. Sex differences in behaviours may skew data (Gruene et al., 2015; Clipperton-Allen et al., 2010) and interpretations (Tronson and Keiser, 2019; Shansky, 2019). It is important for funding agencies and the research community to be aware that beyond incorporation, the specialization of sex differences research is important to propel research forward. This is why organizations such as Organization for the Study of Sex Differences will remain vital to educating the research community to the importance of the science of sex differences in the future.

Furthermore, it is not clear that the mandates are increasing research into both sexes. At CIHR, the mandate goes beyond simple integration of both sexes in research, and requires researchers to incorporate SGBA in their studies. Exploring publicly available data of CIHR grants from 2009–2018 indicate relatively few grants are specifically studying sex differences. Of the project grants (akin to RO1 grants at NIH) from the last two years (2016–2018) awarded in the province of British Columbia, only 2% specified sex differences in their grant descriptions and this percentage did not change (2%) from operating grants from 10 years ago (2009–2015). Thus, while the inclusion of the mandatory questions on sex and gender are laudable and undeniably getting researchers to recognise the importance of study sex and gender, the data do not support a large uptake in grants specifically examining sex and/or gender.

If the impetus behind these initiatives is to improve women's health, grant mechanisms (requests for funding) specifically for women's health would do more to move the dial on women's health. An action that is needed is a groundswell of national and international support, of the kind that fueled funding and research for a moon landing. In the case of biomedical research, one needs to look no further than funding for HIV, as an example of how specific funding initiatives can significantly improve health. The increase in funding for HIV research has been staggering in the last 30 years, increasing from \$2.3B to almost \$35B from 1989 to 2019 in the US (<https://www.kff.org/hiv/aids/slide/federal-hiv-funding-fy-1981-fy-2019/>). Breakthroughs in HIV have been astounding and it no longer carries the death sentence it once did: in 1996 life expectancy of a 20-year-old infected was 30 years, while in 2011 life expectancy was 70 years (Samji et al., 2013). Recently it was announced the 2nd person who previously had HIV and is now living virus free. This would not have been possible without the amount of research on HIV in the last 20 years. We propose that a similar concerted national/international effort is necessary for improving health outcomes necessary for women's or sex-specific health. A surge in investment could potentially make a significant difference in improving the health of women as well as men.

Acknowledgements

The authors acknowledge funding sources that have supported their work. Natural Science and Engineering Research Council of Canada (RGPIN-2018-04301) and Canadian Institutes for Health Research (MOP 142308) to LAMG. National Institutes of Health AG042189 and NS074895 to FS.

References

- Alberich S, Fernández-Sevillano J, González-Ortega I, Usall J, Sáenz M, González-Fraile E, González-Pinto A. A systematic review of sex-based differences in effectiveness and adverse effects of clozapine. *Psychiatry Res.* 2019 10;280:112506. [PubMed: 31401291]
- Becker JB, Koob GF. Sex Differences in Animal Models: Focus on Addiction. *Pharmacol Rev.* 2016 4;68(2):242–63 [PubMed: 26772794]
- Becker JB, Prendergast BJ, Liang JW. Female rats are not more variable than male rats: a meta-analysis of neuroscience studies. *Biol Sex Differ.* 2016 7 26;7:34. doi: 10.1186/s13293-016-0087-5. eCollection 2016. [PubMed: 27468347]
- Beery AK, Zucker I. Sex bias in neuroscience and biomedical research. *Neurosci Biobehav Rev.* 2011 1;35(3):565–72 [PubMed: 20620164]
- Brookmeyer R, Gray S, Kawas C (1998). Projections of Alzheimer's disease in the United States and the public health impact of delaying disease onset. *Am J Public Health* 88: 1337–1342. [PubMed: 9736873]
- Bryant NL, Buchanan RW, Vadar K, Breier A, Rothman M. Gender differences in temporal lobe structures of patients with schizophrenia: a volumetric MRI study. *Am J Psychiatry.* 1999;156(4):603–9. [PubMed: 10200741]
- Buch T, Moos K, Ferreira FM, Fröhlich H, Gebhard C, Tresch A. Benefits of a factorial design focusing on inclusion of female and male animals in one experiment. *J Mol Med (Berl).* 2019 6;97(6):871–877. doi: 10.1007/s00109-019-01774-0. [PubMed: 30980104]
- Cahill L and Hall ED. Is it time to resurrect “lazaroids”? *J Neurosci Res.* 2017 95(1–2): p. 17–20. [PubMed: 27870398]
- Chen PB, Hu RK, Wu YE, Pan L, Huang S, Micevych PE, Hong W. Sexually Dimorphic Control of Parenting Behavior by the Medial Amygdala. *Cell.* 2019 2 21;176(5):1206–1221.e18. [PubMed: 30773317]

- Clipperton Allen AE, Cragg CL, Wood AJ, Pfaff DW, Choleris E. Agonistic behavior in males and females: effects of an estrogen receptor beta agonist in gonadectomized and gonadally intact mice. *Psychoneuroendocrinology*. 2010 8;35(7):1008–22 [PubMed: 20129736]
- Cohen J. 1988 Statistical power analysis for the behavioral sciences. 2nd ed Hillsdale NJ, Lawrence Erlbaum Associates, Publishers.
- Davey DA. Androgens in women before and after the menopause and post bilateral oophorectomy: clinical effects and indications for testosterone therapy. *Womens Health (Lond)*. 2012 7;8(4):437–46. doi: 10.2217/whe.12.27. [PubMed: 22757734]
- De Bellis A, De Angelis G, Fabris E, Cannatà A, Merlo M, Sinagra G. Gender-related differences in heart failure: beyond the “one-size-fits-all” paradigm. *Heart Fail Rev*. 2019 7 2. doi: 10.1007/s10741-019-09824-y.
- Dominoni DM, Borniger JC, Nelson RJ. Light at night, clocks and health: from humans to wild organisms. *Biol Lett*. 2016 2;12(2):20160015. [PubMed: 26888917]
- Durgan DJ, et al., Short communication: ischemia/reperfusion tolerance is time-of-day-dependent: mediation by the cardiomyocyte circadian clock. *Circulation research*, 2010 106(3): p. 546–550. [PubMed: 20007913]
- Ehrhardt S, Appel LJ, Meinert CL. Trends in National Institutes of Health Funding for Clinical Trials Registered in ClinicalTrials.gov. *JAMA*. 2015 12 15;314(23):2566–7 [PubMed: 26670975]
- Eid RS, Gobinath AR, Galea LAM. Sex differences in depression: Insights from clinical and preclinical studies. *Prog Neurobiol*. 2019 5;176:86–102 [PubMed: 30721749]
- Fonken LK, Bedrosian TA, Zhang N, Weil ZM, DeVries AC, Nelson RJ. Dim light at night impairs recovery from global cerebral ischemia. *Exp Neurol*. 2019 7;317:100–109 [PubMed: 30822422]
- Geller SE, Koch AR, Roesch P, Filut A, Hallgren E, Carnes M. The More Things Change, the More They Stay the Same: A Study to Evaluate Compliance With Inclusion and Assessment of Women and Minorities in Randomized Controlled Trials. *Acad Med*. 2018 4;93(4):630–635. doi: 10.1097/ACM.0000000000002027. [PubMed: 29053489]
- Gierisch JM, Coeytaux RR, Urrutia RP, Havrilesky LJ, Moorman PG, Lowery WJ, Dinan M, McBroom AJ, Hasselblad V, Sanders GD, Myers ER. Oral contraceptive use and risk of breast, cervical, colorectal, and endometrial cancers: a systematic review. *Cancer Epidemiol Biomarkers Prev*. 2013 11;22(11):1931–43. doi: 10.1158/1055-9965.EPI-13-0298 [PubMed: 24014598]
- Giustino G, Overbey J, Taylor D, Ailawadi G, Kirkwood K, DeRose J, Gillinov MA, Dagenais F, Mayer ML, Moskowitz A, Bagiella E, Miller M, Grayburn P, Smith PK, Gelijns A, O’Gara P, Acker M, Lala A, Hung J. Sex-Based Differences in Outcomes After Mitral Valve Surgery for Severe Ischemic Mitral Regurgitation: From the Cardiothoracic Surgical Trials Network. *JACC Heart Fail*. 2019 6;7(6):481–490. doi: 10.1016/j.jchf.2019.03.001. [PubMed: 31146872]
- Goel N, Workman JL, Lee TT, Innala L, Viau V. Sex differences in the HPA axis. *Compr Physiol*. 2014 7;4(3):1121–55 [PubMed: 24944032]
- Golden LC, Voskuhl R. The importance of studying sex differences in disease: The example of multiple sclerosis. *J Neurosci Res*. 2017 1 2;95(1–2):633–643. doi: 10.1002/jnr.23955. [PubMed: 27870415]
- Greenblatt DJ, Harmatz JS, von Moltke LL, Wright CE, Durol AL, Harrel-Joseph LM, Shader RI. Comparative kinetics and response to the benzodiazepine agonists triazolam and zolpidem: evaluation of sex-dependent differences. *J Pharmacol Exp Ther*. 2000 5;293(2):435–43. [PubMed: 10773013]
- Greenblatt DJ, Harmatz JS, Roth T, Singh NN, Moline ML, Harris SC, Kapil RP. Comparison of pharmacokinetic profiles of zolpidem buffered sublingual tablet and zolpidem oral immediate-release tablet: results from a single-center, single-dose, randomized, open-label crossover study in healthy adults. *Clin Ther*. 2013; 35(5):604–11. doi: 10.1016/j.clinthera.2013.03.007. [PubMed: 23541711]
- Gruene TM, Flick K, Stefano A, Shea SD, Shansky RM. Sexually divergent expression of active and passive conditioned fear responses in rats. *Elife*. 2015 11 14;4 pii: e11352. doi: 10.7554/eLife.11352. [PubMed: 26568307]
- Gutiérrez-Lobos K, Scherer M, Anderer P, Katschnig H (2002). The influence of age on the female/male ratio of treated incidence rates in depression. *BMC Psychiatry* 2: 3. [PubMed: 11860609]

- Haack S, Seeringer A, Thürmann PA, Becker T, Kirchheiner J. Sex-specific differences in side effects of psychotropic drugs: genes or gender? *Pharmacogenomics*. 2009 9;10(9):1511–26. doi: 10.2217/pgs.09.102 [PubMed: 19761372]
- Häfner H, Riecher-Rössler A, Maurer K, Fätkenheuer B, Löffler W. First onset and early symptomatology of schizophrenia. A chapter of epidemiological and neurobiological research into age and sex differences. *Eur Arch Psychiatry Clin Neurosci*. 1992;242(2–3):109–18. [PubMed: 1486099]
- Harden KP, Wrzus C, Luong G, Grotzinger A, Bajbouj M, Rauters A, Wagner GG, Riediger M. Diurnal coupling between testosterone and cortisol from adolescence to older adulthood. *Psychoneuroendocrinology*. 2016 Nov;73:79–90
- Harman SM, Metter EJ, Tobin JD, Pearson J, Blackman MR; Baltimore Longitudinal Study of Aging. Longitudinal effects of aging on serum total and free testosterone levels in healthy men. Baltimore Longitudinal Study of Aging. *J Clin Endocrinol Metab*. 2001 2;86(2):724–31 [PubMed: 11158037]
- Hinkula M, Kauppila A, Nyyhää S, Pukkala E. Cause-specific mortality of grand multiparous women in Finland. *Am J Epidemiol*. 2006 2 15;163(4):367–73. Epub 2005 Dec 21. [PubMed: 16371514]
- Irvine K, Laws KR, Gale TM, Kondel TK (2012). Greater cognitive deterioration in women than men with Alzheimer's disease: a meta analysis. *J Clin Exp Neuropsychol*. 34: 989–98 [PubMed: 22913619]
- Jang H, Bae JB, Dardiotis E, Scarmeas N, Sachdev PS, Lipnicki DM, Han JW, Kim TH, Kwak KP, Kim BJ, Kim SG, Kim JL, Moon SW, Park JH, Ryu SH, Youn JC, Lee DY, Lee DW, Lee SB, Lee JJ, Jhoo JH, Yannakoulia M, Kosmidis MH, Hadjigeorgiou GM, Sakka P, Kim KW Differential effects of completed and incomplete pregnancies on the risk of Alzheimer disease. *Neurology*, 2018 91: e643–e651. [PubMed: 30021919]
- Kassell NF, et al., Randomized, double-blind, vehicle-controlled trial of tirilazad mesylate in patients with aneurysmal subarachnoid hemorrhage: a cooperative study in Europe, Australia, and New Zealand. *J Neurosurg*, 1996 84(2): p. 221–8. [PubMed: 8592224]
- Keyes KM, Cheslack-Postava K, Westhoff C, Heim CM, Haloossim M, Walsh K, Koenen K. Association of hormonal contraceptive use with reduced levels of depressive symptoms: a national study of sexually active women in the United States. *Am J Epidemiol*. 2013 11 1;178(9):1378–88. [PubMed: 24043440]
- Kosmidis MH, Hadjigeorgiou GM, Sakka P, Kim KW. Differential effects of completed and incomplete pregnancies on the risk of Alzheimer disease. *Neurology* Aug 2018, 91 (7) e643–e651; DOI: 10.1212/WNL.0000000000006000
- Kessler RC, Bromet EJ, 2013 The epidemiology of depression across cultures. *Annu. Rev. Public Health* 34, 119–38. doi:10.1146/annurev-publhealth-031912-114409 [PubMed: 23514317]
- Kim JH, Lee SJ. Parity and increased risk of insulin resistance in postmenopausal women: the 2010 Korean National Health and Nutrition Examination Survey. *Menopause*. 2017 7;24(7):832–837. doi: 10.1097/GME.0000000000000846. [PubMed: 28291026]
- Labots G, Jones A, de Visser SJ, Rissmann R, Burggraaf J. Gender differences in clinical registration trials: is there a real problem? *Br J Clin Pharmacol*. 2018 4;84(4):700–707. doi: 10.1111/bcp.13497. [PubMed: 29293280]
- Lang XE, Zhu D, Zhang G, Du X, Jia Q, Yin G, Chen D, Xiu M, Cao B, Wang L, Li X, Soares JC, Zhang XY7. Sex difference in association of symptoms and white matter deficits in first-episode and drug-naïve schizophrenia. *Transl Psychiatry*. 2018 12 18;8(1):281. doi: 10.1038/s41398-018-0346-9. [PubMed: 30563964]
- Leveille SG, Resnick HE, Balfour J. Gender differences in disability: evidence and underlying reasons. *Aging (Milano)*. 2000 4;12(2):106–12 [PubMed: 10902052]
- Mahmood SS, et al., The Framingham Heart Study and the epidemiology of cardiovascular disease: a historical perspective. *Lancet (London, England)*, 2014 383(9921): p. 999–1008.
- Mattina GF, Steiner M. The need for inclusion of sex and age of onset variables in genetic association studies of obsessive-compulsive disorder: Overview. *Prog Neuropsychopharmacol Biol Psychiatry*. 2016 6 3;67:107–16. [PubMed: 26827635]

- Mensah GA, et al., Decline in Cardiovascular Mortality: Possible Causes and Implications. *Circulation research*, 2017 120(2): p. 366–380. [PubMed: 28104770]
- Mersha TB, Martin LJ, Biagini Myers JM, Kovacic MB, He H, Lindsey M, Sivaprasad U, Chen W, Khurana Hershey GK. Genomic architecture of asthma differs by sex. *Genomics*. 2015 7;106(1):15–22. doi: 10.1016/j.ygeno.2015.03.003. [PubMed: 25817197]
- McPherson S, Back C, Buckwalter JG, Cummings JL (1999). Gender-related cognitive deficits in Alzheimer's disease. *Int Psychogeriatrics* 11: 117–122.
- Prendergast BJ, Onishi KG, Zucker I. Female mice liberated for inclusion in neuroscience and biomedical research. *Neurosci Biobehav Rev*. 2014;40:1–5 [PubMed: 24456941]
- Previtato M, Osto E, Kerkhof PLM, Parry G, Tona F (2018) Heart Transplantation Survival and Sex-Related Differences In: Kerkhof P, Miller V (eds) *Sex-Specific Analysis of Cardiovascular Function*. *Advances in Experimental Medicine and Biology*, vol 1065 Springer, Cham
- Reddy DS. The neuroendocrine basis of sex differences in epilepsy. *Pharmacol Biochem Behav*. 2017 1;152:97–104. doi: 10.1016/j.pbb.2016.07.002. Epub 2016 Jul 14. [PubMed: 27424276]
- Rilling JK, Demarco AC, Hackett PD, Chen X, Gautam P, Stair S, Haroon E, Thompson R, Ditzen B, Patel R, Pagnoni G. 2014 Sex differences in the neural and behavioral response to intranasal oxytocin and vasopressin during human social interaction. *Psychoneuroendocrinology*. 1;39:237–48 [PubMed: 24157401]
- Shansky RM. Are hormones a “female problem” for animal research? *Science*. 2019 5 31;364(6443):825–826. doi: 10.1126/science.aaw7570. [PubMed: 31147505]
- Shansky RM, Woolley CS. Considering Sex as a Biological Variable Will Be Valuable for Neuroscience Research. *J Neurosci*. 2016 11 23;36(47):11817–11822. [PubMed: 27881768]
- Skovlund CW, Mørch LS, Kessing LV, Lidegaard Ø. Association of Hormonal Contraception With Depression. *JAMA Psychiatry*. 2016 11 1;73(11):1154–1162. doi: 10.1001/jamapsychiatry.2016.2387. [PubMed: 27680324]
- Sohrabji F, Welsh CJ, and Reddy DS, Chapter 12 - Sex Differences in Neurological Diseases, in *Sex Differences in the Central Nervous System*, Shansky RM, Editor. 2016, Academic Press: San Diego p. 297–323.
- Sohrabji F, Park MJ, and Mahnke AH, Sex differences in stroke therapies. *J Neurosci Res*, 2017 95(1–2): p. 681–691. [PubMed: 27870437]
- Sorge RE, LaCroix-Fralish ML, Tuttle AH, Sotocinal SG, Austin JS, Ritchie J, Chanda ML, Graham AC, Topham L, Beggs S, Salter MW, Mogil JS. Spinal cord Toll-like receptor 4 mediates inflammatory and neuropathic hypersensitivity in male but not female mice. *J Neurosci*. 2011 10 26;31(43):15450–4. doi: 10.1523/JNEUROSCI.3859-11.2011. [PubMed: 22031891]
- Sorge RE, Mapplebeck JC, Rosen S, Beggs S, Taves S, Alexander JK, Martin LJ, Austin JS, Sotocinal SG, Chen D, Yang M, Shi XQ, Huang H6, Pilon NJ, Bilan PJ, Tu Y, Klip A, Ji RR, Zhang J, Salter MW, Mogil JS. Different immune cells mediate mechanical pain hypersensitivity in male and female mice. *Nat Neurosci*. 2015 8;18(8):1081–3. doi: 10.1038/nn.4053. Epub 2015 Jun 29. [PubMed: 26120961]
- Suárez-Barrientos A., et al., Circadian variations of infarct size in acute myocardial infarction. *Heart*, 2011 97(12): p. 970. [PubMed: 21525526]
- Tackett JL, Reardon KW, Herzhoff K, Page-Gould E, Harden KP, Josephs RA. Estradiol and cortisol interactions in youth externalizing psychopathology. *Psychoneuroendocrinology*. 2015 5;55:146–53. [PubMed: 25765756]
- Tronson NC, Keiser AA. A Dynamic Memory Systems Framework for Sex Differences in Fear Memory. *Trends Neurosci*. 2019 10;42(10):680–692. [PubMed: 31473031]
- Viau V. Functional cross-talk between the hypothalamic-pituitary-gonadal and -adrenal axes. *J Neuroendocrinol*. 2002 6;14(6):506–13 [PubMed: 12047726]
- Wada H, Miyauchi K, Daida H. Gender differences in the clinical features and outcomes of patients with coronary artery disease. *Expert Rev Cardiovasc Ther*. 2019 2;17(2):127–133 [PubMed: 30569774]
- Westergaard David, Moseley Pope, Karuna Freja Sørup Hemmingsen, Baldi Pierre, Brunak Søren. Population-wide analysis of differences in disease progression patterns in men and women. *Nature Communications*, 2019; 10 (1) DOI: 10.1038/s41467-019-08475-9

- Will TR, Proaño SB, Thomas AM, Kunz LM, Thompson KC, Ginnari LA, Jones CH, Lucas S-C, Reavis EM, Dorris DM, Meitzen J. Problems and Progress regarding Sex Bias and Omission in Neuroscience Research. *eNeuro* 3 11 2017, 4 (6) ENEURO.0278–17.2017; DOI: 10.1523/ENEURO.0278-17.2017
- Williamson CM, Lee W, Romeo RD, Curley JP. Social context-dependent relationships between mouse dominance rank and plasma hormone levels. *Physiol Behav.* 2017 3 15;171:110–119. doi: 10.1016/j.physbeh.2016.12.038. Epub 2017 Jan 6. [PubMed: 28065723]
- Woitowich NC, Woodruff TK. Implementation of the NIH Sex-Inclusion Policy: Attitudes and Opinions of Study Section Members. *J Womens Health (Larchmt).* 2019 1;28(1):9–16. [PubMed: 30539677]
- Woitowich NC, Woodruff TK. Opinion: Research community needs to better appreciate the value of sex-based research. *Proc Natl Acad Sci U S A.* 2019 4 9;116(15):7154–7156. doi: 10.1073/pnas.1903586116. [PubMed: 30971497]
- Woitowich NC, Woodruff TK. No female mice or cells? NIH reviewers still might score grant OK. *Nature.* 2019c 1; 565(7737):25
- Wright DW, Yeatts SD, Silbergleit R, Palesch YY, Hertzberg VS, Frankel M, Goldstein FC, Caveney AF, Howlett-Smith H, Bengelink EM, Manley GT, Merck LH, Janis LS, Barsan WG; NETT Investigators. Very early administration of progesterone for acute traumatic brain injury. *N Engl J Med.* 2014 12 25;371(26):2457–66. doi: 10.1056/NEJMoa1404304. [PubMed: 25493974]
- Young H, Oreve MJ, Speranza M. Clinical characteristics and problems diagnosing autism spectrum disorder in girls. *Arch Pediatr.* 2018 8;25(6):399–40 [PubMed: 30143373]
- Zeng Z, Surewaard BGJ, Wong CHY, Guettler C, Petri B, Burkhard R, Wyss M, Le Moual H, Devinney R, Thompson GC, Blackwood J, Joffe AR, McCoy KD, Jenne CN, Kubes P. Sex-hormone-driven innate antibodies protect females and infants against EPEC infection. *Nat Immunol.* 2018 10;19(10):1100–1111. doi: 10.1038/s41590-018-0211-2. [PubMed: 30250184]

- Sex differences are important to consider in biomedical research
- Appropriate sample sizes are needed for adequate power to detect sex differences
- Implementation issues for sex as a biological variable need to be considered
- Integration and specialisation for the study of sex differences is imperative
- Fluctuating steroid hormones are noted in both sexes that influence outcome variables

Figure 1.

Estimated power from simulations (500 replicates) using SD = 2, mean in male control = 10, mean in female control = 8, mean in male treatment = 15, and varying means in female treatment = 1) 3, 2) 8, 3) 10.5, and 4) 11.75.

Reviewer Guidance to Evaluate Sex as a Biological Variable (SABV)

Main points

- NIH expects that sex as a biological variable will be factored into research designs, analyses, and reporting in vertebrate animal and human studies.
- Strong justification from the scientific literature, preliminary data, or other relevant considerations must be provided for applications proposing to study only one sex.
- This decision tree is meant to be used as a guide, but does not encompass the entire policy. See [NOT-OD-15-102](#) for more information.

Figure 2.

Suggested changes (A,B) to National Institute of Health's Reviewer Guidance to evaluate Sex as a Biological Variable. Some reviewers may need reminders that single sex research is not limited to ovarian or testicular research.

Proposed changes to flowchart:

A: Is the proposal powered for detection of sex differences?

B: For example, is it for sex-specific reason: ie oral contraceptive use, menstrual cycle, prostate cancer, menopause, etc.