Female Scientist and Biomedicine*

Daneal O'Habib
University of Waterloo
d3ohabib@edu.uwaterloo.ca

December 3, 2019

Abstract

Are female scientists more likely to create medical research aimed specifically at women? In this paper, I investigate whether the gender of the scientist shapes the nature of their scientific research. Analyzing data on the universe of biomedical research, I find that papers with female scientists are more likely to focus on conditions that only impact women. Consistent with the idea of female researchers choosing to innovate for women, I find stronger effects when women are in positions of authority on scientific publications. Women-led research papers are 40 percent more likely to focus on female specific diseases and conditions. This result suggests that increasing gender equality in STEM may have broader implications for female health outcomes than previously suspected.

JEL classification: XXX, YYY.

Keywords: Economics of Innovation, Gender and Economics, Women and STEM

*I wish to thank Joel Blit for his supervision and encouragement throughout this project. I also want to thank Mikko Packalen for exposing me to the Medline data.

Introduction

It is taken for granted that there are diseases that only effect women. Despite the obvious implications of this banal observation, the medical profession has been slow to allow women to enter and conduct research, effectively leaving men the study these highly gendered diseases. The results have been predictably one-sided and a structural neglect of female health has been documented in all areas of research.

For instance, even up until the late 1980's research into breast cancer – which has among the highest incident rates of all cancers—lagged behind cancers with lower incident rates (Lukong, 2017). In medical trials, male subjects were treated as the benchmark model, and female subjects were treated as an afterthought due to "hormonal responses" complicating research results (Klein, 2019). Researchers assumed that findings from studies on male subjects would apply to females as well. But this has proven to be false, and women's health as suffered as direct consequences. Funding for research also shows a pervasive gender bias. For many years, research in biomedicine overlooked diseases that disproportionately affected women, with male specific diseases receiving a much larger share of funding and research attention (Barlow, 2018). The history of medicine is littered with embarrassing oversights surrounding female specific maladies, but it seems that recent history also paints a bleak picture.

This structural neglect of female research has associated with the historic underrepresentation of women in science (Barlow, 2018; Kahn and Ginther, 2017; Nielsen et al., 2017a). Indeed, research has shown that a scientist pre-existing preference shapes their current research agenda and priorities (Azoulay et al., 2017). But many disciplines in science are starting to achieve gender parity, and research related to women's health has become more common as female representation in biomedicine has grown (Office of Research on Women's Health, 2016). Since the sexes differ in their exposure to specific diseases, and women in general may be attuned to female specific maladies, this fact should not be surprising.

This background information gestures at a phenomena that is received little attention in the economics of innovation: that the demographics of inventors matters for the what is invented. Indeed, major health innovations spearhead by male scientists may not have been possible without female insight. For example, Gregory Pincus is credited with inventing the birth control pill in the 1960's. But his interest in this topic began after feminist Margaret Sanger urged him to work on the problem at a dinner in 1951 (Little, 2014). Furthermore, funding for research on the birth control pill remained elusive until Katharine Dexter McCormick, the 2nd woman to graduate from MIT, funded the research (Little, 2014). Counterfactuals are always difficult, but ones wonders how much faster female reproductive health could have advanced if women had a great role in science decades earlier.

It is plausible that the lack of attention paid to female specific disease is associated with female underrepresentation in biomedicine. But empirical evidence demonstrating that female scientists are more likely to create science focusing women's health is non-existent. I wish to fill this knowledge gap and test whether female scientists, in particular female scientists working in biomedicine, are more likely to target diseases that disproportionately affect women. This paper is structured as follows. First, I conduct a literature reviewing on topics that are may to relevant to our discussion. I use this literature review to narrow down the scope of the research and formulate a clear hypothesis that can be tested. I proceed with sketch out my empirical strategy and data sources. Lastly, I present the results and outline the implications of my findings.

Literature Review

Gender and Innovation

It should be noted that economists that study innovation have not ignored gender; rather, the existing research studies the gender gaps in labour market outcomes and productivity measures between scientist and inventors. For instance, Bell et al. (2019), studying inventors in the United States, found that women are more likely to file patents when they have been exposed to female inventors in their childhood. More to the point, these female inventors are much more likely to innovate in the specific subfield of their childhood role model. This finding underscores the importance of role models and network effects in determining not only who becomes an inventor and what gets invented. There may be an analogue within the biomedical sciences: if women are indeed more likely to create biomedical science for other women, a lack exposure to female leadership in STEM professions may be costing us research output aimed at women (Kahn and Ginther, 2017).

But men still produce patents at a much higher rate, and productivity differences between male and female researchers are manifest in even their late-stage career outputs. Murray and Graham (2007) have argued that is because men have more experience with the process of commercialising, whereas women are still relative new-comers in this field. To make matters worse, experimental evidence has shown that women are less likely to be granted a patent, despite submitting equivalent inventions for review (Jensen et al., 2018).

Lerchenmueller and Sorenson (2018) have documented the female under-represented in management position within research intuitions. This is important because subordinates have less control over their research direction, and desire to create female specific research may stifled due to lack of research freedom.

It also may also be the case that women simply choose to sort into research areas that are not at the cutting edge - and therefore women may be less likely to contribute to ground breaking research. Although this is still an open question (Murray and Graham, 2007). However, it is established that gender bias in STEM professions stifles the research productivity of women (Kahn and Ginther, 2017; Lerchenmueller and Sorenson, 2018).

This line of research, while interesting and deepens our understanding of how gender intersects with the economics of innovation, leaves us with the knowledge gap I outlined above. Our current understanding of innovations only considers one dimension: the rate of innovation. Koning et al. (2019), analyzing biomedical patents over the last several decades, have studied the relationship between the demography of inventors and what actually gets invented. Their results show that patents filed by women are more likely to target diseases and conditions that impact women.

If it is shown that demography of scientist matters for what type of research is conducted – i.e., the increase in female scientist has increased our understanding of diseases and conditions that disproportionately women—it forces us to consider not only the rate of innovation, but also the nature what is invented and who is likely to benefit. It must be said, though, that for most areas of scientific innovation this distinction serves no practical purpose. It is only in biomedicine where it possible to do "gendered" research as there can be an asymmetry in who benefits from a particular research agenda (Nielsen et al., 2017a).

Discrimination and Innovation

Since research usually happens in a labs, the institutional context in which science is created deserves our attentions. Despite the fact the female representation in science is increasing, all new scientist enter their labs and PhD programs at the bottom, and therefore have to deal with norms and structures that were created before they arrived. If female health has a history of being neglected, it may be the case that women simply adopt these norms and work on problems proposed by more senior researchers.

There is a parallel to draw between my research question and the line of research studying the effect of female representation and female promotions in the private sector. Srivastava and Sherman (2015) found that women evaluate other women just as harshly as other their male counter-parts, and only a minority of women mentor and advocate for their female subordinates. The story put forth in this paper is that women develop a "cogs in the machine" mentality and simply take on the norms that were present before they arrived. Although other research has shown that organizational practices change only when women reach a certain critical mass in leadership position (Konrad et al., 2008; Torchia et al., 2011). This

difference is has been linked to selection bias in organizational structures. If an institution has a history of excluding and downplaying women's preferences, the women that excel in these structures are may be those that best conform to prevailing norms and attitudes. Even when women reject such biased attitudes and norms, they have to operated in institutions where evaluation criteria favor men. Changing such practices, then, requires the cooperation and will of multiple individuals that have the authority to make such changes (Konrad et al., 2008; Torchia et al., 2011).

The same points can hold in biomedical research. Although female representation in the sciences has increased, all new scientists enter organizations at the lower levels (Holman et al., 2018). Recent cohorts of female scientists are younger and had less job tenure than their male counter-parts for the last generation; it has been argued that this reflects the historic exclusion women have faced from medical research institutions (Lerchenmueller and Sorenson, 2018; Long and Fox, 1995). This is relevant to my question because younger scientists have less freedom to set their own research agendas. Therefore, even if we assume that women are inclined to work on female conditions, it is possible that women simply have few opportunities to work on diseases and conditions affecting women. At the very least, this suggests that women would have more influence on the types of inventions when they hold more leadership research positions.

Hypothesis

Based on the research reviewed above, I state three hypotheses. The first is that there is an association between a scientists gender and the gender focus of their research:

Hypothesis (1): *Papers with a female scientist are more likely to target diseases or conditions that disproportionately affect women*

Scientific research takes place in institutions where women have historically been under-represented, so the relationship posited in hypothesis (1) could be stronger when women are in leadership positions that gives them more freedom. This assumes that women have latent desires to pursue female specific research that can only be expressed once they can actually control their own research agenda.

Hypothesis (2): *The effect associated with hypothesis 1 is stronger when the women are in leadership positions in scientific publications*

As mentioned at in the literature review, younger scientists don't have control over their research agenda, and if women work in labs that are biased, they will simply reproduce biased scientific output. In this respect, research labs may parallel other institutions, where female managers make decisions similar to their male counter-parts. Based on the research cited in the literature review, the magnitude or even

direction of the effect is not obvious.

Hypothesis (3): *Greater gender-equality is associated with female scientists working on diseases or conditions that disproportionately affect women*

The background information supports the idea social norms and biases pushed male scientist to overlook female-specifics diseases. These biases can have an insidious impact on all stages of the research process - e.g., what research topics get funded - so greater gender-equality could be mediating factor in driving female specific biomedical research. But even here the effect size remains unclear. For example, in countries with high gender equality, males may now be more attenuated female specific diseases, and effect size posited in hypothesis (1) may not be significant.

In the next section, I map out an empirical strategy that allows me to tests the conjectures listed above.

Data and Empirical Strategy

My primary data source is the Medline database, a database of approximately 22 million journal articles that covers nearly every biomedical article published since 1946. To classify research papers by the disease or condition they target, I draw on the National Library of Medicine's Medical Text Indexer (Aronson et al., 2004), which generates Medical Subject Headings (MeSH) for a given body of text. This Medical Text Indexer uses the National Library of Medicine's proprietary natural language processing algorithm to curate the biomedical literature. My analysis rests on the link between the text biomedical text (from the abstracts and titles), and the MeSH produced by the MTI output (Mork et al., 2013). See appendix C for more detail.

Female Focus of Biomedical Research

To measure whether research papers differentially impact males or females, I use the National Library of Medicine's Medical Text Indexer (MTI) to generate Medical Subject Headings (MeSH) terms from the abstract and title of each PubMed ID[1]. Within the MeSH ontology, the MeSH "Male" applied to any research covering "male organs, diseases, physiologic processes and genetics". Conversely, the MeSH "Female" applies to any research covering "female organs, diseases, physiologic processes and genetics". These terms give us a binary indicator whether the paper is more or less likely to be male-or female-focused. If a paper returns (1) a "female" check tag, and (2) no "male" check tag, I create a dependant variable

[1]see Appendix C for why using manually indexed MeSH leads to measurement error

called "Female MeSH". Of course, we do the opposite and get a "Male MeSH" variable, allowing us to capture highly gendered research papers.

I should note that the MTI takes a liberal approach to assigning tags and may return medical subject headings that are only loosely related to the topic of the paper. I discard all tags with low scores, leaving me with an maximum of 25 tags on each paper. This helps correct for the change in scope of biomedical research over the years, with the titles and abstracts having increased both in the number of words, and the topics covered.

My second dependant variable focuses on highly gendered diseases differences. The MeSH terms generated are organized into a parent-child tree, and at the top level are general categories, e.g., Anatomy [A], Organisms [B] and Diseases [C]. Within "diseases" (branch C) there are more focused disease classes:

- Nervous System Diseases [C10]
- Eye Diseases [C11]
- Male Urogenital Diseases [C12]
- Female Urogenital Diseases and Pregnancy Complications [C13]

I focus on branch [C12] and [C13]. The urologic diseases are the only thing I found that is (1) highly gendered, and (2) fits the MTI classification system so we can code binary gender indicators for specialized research areas.[2] Any paper that matches to [C12] — or any term further down the knowledge tree (e.g., Azoospermia [C12.294.365.700.380]) — can be coded as a male disease; and paper that matches [C13] and below can be coded as a female disease. This gives me a fine-grained measure and adds credibility to the previous variable. While the gendered MeSH dependent variable has some limitations, this variable unequivocally captures highly gendered research and even provides a male/female within a specific disease category.

It should be noted that the "male" and "female" MeSH mentioned doesn't have any tree code; these tags only have unique concept identifiers. Within the MeSH ontology, these coarse unique concept identifiers sit atop more granular MeSH which always has a tree code. With respect to text analysis of PubMed citations, the "female" MeSH will always accompanies a more granular MeSH and tree code that captures the specific female disease under study. A very specific urogenital disease studied within the [C13] branch may be research into invasive hydatidiform moles [C13.703.720.949.416.875.500].

[2]please see Appendix B for a visual of the MeSH ontology

Limitations of the Medical Text Indexer

The MeSH terms generated by the MTI offers a reasonable approximation about whether a piece of medical text is related to men or women. Like any measure, these male and female tags have some limitations. While it true that the MTI is starting to be used in conjunction with manual indexing done by humans, the main purpose of the entire endeavor is to help researchers find related research on the basis of medical subject headings. Therefore, it may conflate research that is specifically aimed at helping women - e.g., a new treatment for breast cancer - with research that helps women indirectly - e.g., a paper that presents a case study of drug overdose where the subject happens to be female.[3] We can clearly see that the latter topic has a tenuous relationship to the measure of interest; therefore, I create a robustness check using a dictionary matching approach. I adopt a list of medical terms that only impact men or women, and search the title and abstract of each citation under study. I match title and abstract text to terms in the dictionary. All papers that match to a male term in the dictionary are coded as "Male" and all papers that match to a term in the female dictionary are flagged as "Female".[4] Of course, this will miss a lot since science has gotten very specialized, and the keywords I may be too general to capture everything. But it is helpful to have has a robustness check that completely avoids the MTI algorithm.

Gender

Gender was inferred by using genderize.io, an API that curates name-gender associations by crawling the web. This API has over 80 countries in its database, ensuring that I can accurately infer ethnic first names for author in the Medline database. Country affiliation of authors was also specified before querying the database, ensuring I avoid any idiosyncratic country specific name-gender associations. For example, the name Kim is a female name in the Canada, but it is a common male name in Denmark.

The proportion of individuals with any specific name in the database corresponds to a "gender score" - i.e., 10% of people women in America may have the name Chris. I address this ambiguity by removing all names that don't match to a gender with at least a 95% gender score. This step preserved approximately 93% of the data set. The average gender score of all excluded names was 0.47 and the median was 0.49, ensuring that our measurement is unbiased.

[3] see Appendix C for more on this topic

[4] please see Appendix D for the list of keywords

Female Leadership

The Medline database includes author position which I used as a proxy for leadership on a paper. Broadly speaking, the last author on a paper sets the research agenda for the lab, and therefore has the most influence on the topics studied. The first author is responsible for carrying out research project and they are generally the most knowledgeable on the specific topic being studied in the paper. I flag all papers with female first and last author and create binary inductors to use as proxies for leadership in scientific output.

Most papers will include a supporting cast that is not directly in leadership roles. I subset the data for all journals with at least three authors, and flag all papers with at least one female name that is not the first or last author. This variable is labelled "female contributor" and their impact on the nature of the research being produced is hypothesized to be small.

Lastly, I create two more measure of female leaderships on papers: (1) "any female" - a binary indictor if any female is present on the paper, irrespective of position, and (2) "share of females" - I take the number of women on the citation and divided that by the total team size.

Gender Equality

I use the United Nations Human Development Report to obtain country level metrics on gender equality (Programme, 2016). The specific metric used is called the Gender Inequality Index, which captures female empowerment across three domains: health empowerment, and economics status. The variables used in the index are maternal morality ratio, share of women in parliament seats, and labour force participation rate of women aged 15 and over. I simply group by country and take the mean value of the index from 2000-2012. I code all countries that are one standard deviation below the mean as "high equality" and the rest are coded as "low equality". Note that the measure is between 0 and 1, with 0 being highly gender-egalitarian countries.

Research Scope

Under more favourable circumstances, I would have tried inferred gender with all data that is available. Regrettably, Medline database only has data author forenames for citations after 2001, so I had to truncate my analysis for years after 2002. Previous research using the Medline author data made similar concessions and truncated their analysis to years after 2002 (Holman et al., 2018; Nielsen et al., 2017b).

The project was further narrowed in scope due to large volumes of data to be processed by the MTI

batch processing tool. There are over 11 million citations in PubMed from 2002-2016 alone, and processing everything was not practical given the time constraints. Even after testing the removal stop-words from the title and abstract text, the processing times remained intractable. I subset the data from 2002-2012 effectivity leaving me with approximately four million citations as a cross section. I randomly sample observations from this cross section and tried to process as many citations as I could. Random sampling ensures that we can get a unbiased estimate within this subsample. But all results should be interpreted with caution and they may not hold once data from 2012 onward is added.

Controls

In all regression specifications, I include a battery of fixed effects to adjust for unobserved heterogeneity across time, research categories, and countries. In the literature review, I raise the possibility that male and female scientist select into different research areas. To ameliorate this bias, I include research category fixed effects for all subdisciplines in the sample. Including years fixed effects helps to disentangle any effect of greater female representation from the secular increase both in women scientist and the number of female-focused researcher papers. I raise the possibility that gender bias impacts the degree to which women are free to pursue research that is overlooked by men. I include country fixed effects to adjust for gender bias and other unobserved factors across countries.

Summary Statistics

Table one provides summary statistics of the variables described above.

Table 1: Summary Statistics

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Number of MeSH Terms on Paper	1,437,097	13.46	5.34	1	9	17	25
Year	1,437,097	2,007.71	3.11	2,002	2,005	2,010	2,012
Team Size	1,437,097	5.17	9.86	1	3	7	3,058
Male MeSH	1,437,097	0.11	0.31	0	0	0	1
Female MeSH	1,437,097	0.10	0.30	0	0	0	1
Female Urogenital Diseases Treecode	1,437,097	0.02	0.14	0	0	0	1
Male Urogenital Diseases Treecode	1,437,097	0.01	0.10	0	0	0	1
Any Female Author on Paper	1,437,097	0.62	0.48	0	0	1	1
Female Contributor*	1,147,106	0.05	0.23	0.00	0.00	0.00	1.00
Female First Author	1,437,097	0.29	0.45	0	0	1	1
Female Last Author	1,437,097	0.19	0.39	0	0	0	1
Proportion Female on Paper	1,437,097	0.26	0.27	0.00	0.00	0.43	1.00

*only papers with at least three authors

Results

Are female scientists are more likely to create science focusing women's health? I regress whether a PubMed citations matches to a "Female" medical subject heading on whether the paper has any female scientist working on it. A direct analogue can be captured by testing the impact of female scientist on "Male" medical subject headings. From Table 2, it appears that women are more likely to create female specific science. This result holds under a battery of controls/fixed effects and the effect size is economically significant. Given that the Female MesH variable has a base rate of 10 percentage points, this translates into a $(10 + 2.3)/10 = 23\%$ increase, which is substantial. The same percentage increase with respect to female specific urogenital diseases is $(2+0.5)/2 = 25\%$ - again an economically meaningful effect. Note that the effect size in the male analogue - both the Male MeSH and Male Urogenital Diseases - is much smaller compared to the female focused dependant variables. Since this is a difficult problem to operationalize, Table 3 tests the same idea but with the proportion of female scientist on a paper as the independent variable. The results parallel those found in Table 2, suggesting that hypothesis (1) bourn out by the data.

The independent variables used in Table 2 and 3 are relatively opaque and not very informative on their own. Tables 4 and 5 offer more clarity and test if greater female authority in scientific papers is associated with increased attention to conditions that affect women. These models demonstrate that the impact of female scientist in non-lead positions is negligible, whereas papers with female first authors produce an economically and statistically significant effect. More to the point, when women have the greatest authority - i.e., when they are the last author on a paper - the effect size seems to be the strongest. Indeed, female last authors are associated with a $(10+3.9)/10 = 39\%$ increase in female focused science, as measured by the Female MeSH variable. This effect jumps to $(2 + 0.9)/2 = 45\%$ increase as measured by Female urogenital diseases and pregnancy complications. This result seems to suggest that hypothesis (2) is bourn out by the data, and greater female authority in scientific papers is associated with greater attention to conditions that only impact women.

With respect the hypothesis (3), I take the "high gender equality" dummy variable and interact that with female last authors. Once again, it seems that the hypothesis borne out by the data, and women are more likely to create female focused science under conditions of greater female equality. Furthermore, I interact the female first author variable with the female last author variable to test if there is something interesting between this interplay that is driving the effect. All coefficients remain significant, both economically and statistically.

Nielsen et al. (2017b) argued that female scientists are more likely to conduct gender analysis - i.e., compare outcomes between male and female rats. This result goes further and argues that women are much more likely to work on diseases and conditions that impact women. The qualifications in Appendix C not withstanding, this is the first paper is first to try and estimate such a relationship in scientific output.

Tables

Table 2: Female Scientist and Biomedicine

	Measures of Gendered Research			
	Female MeSH	Male MeSH	Female Urogenital Disease	Male Urogenital Disease
	(1)	(2)	(3)	(4)
Any Female Author	0.023***	−0.004***	0.005***	−0.002***
	(0.001)	(0.001)	(0.0003)	(0.0002)
Year FE	Yes	Yes	Yes	Yes
Research Category FE	Yes	Yes	Yes	Yes
Team Size Control	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Observations	1,437,097	1,437,097	1,437,097	1,437,097
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 3: Female Scientist and Biomedicine

	Measures of Gendered Research			
	Female MeSH	Male MeSH	Female Urogenital Disease	Male Urogenital Disease
	(1)	(2)	(3)	(4)
Proportion Female Authors	0.060***	−0.016***	0.011***	−0.005***
	(0.001)	(0.001)	(0.0005)	(0.0003)
Year FE	Yes	Yes	Yes	Yes
Research Category FE	Yes	Yes	Yes	Yes
Team Size Control	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Observations	1,437,097	1,437,097	1,437,097	1,437,097
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 4: Stronger Effect Size with Greater Female Leadership

	Gendered Research Categorized by MeSH					
	Female Only MeSH			Male Only Mesh		
	(1)	(2)	(3)	(4)	(5)	(6)
Fem Contributor	0.009***			−0.011***		
	(0.001)			(0.001)		
Fem. First Author		0.024***			−0.006***	
		(0.001)			(0.001)	
Fem. Last Author			0.039***			−0.008***
			(0.001)			(0.001)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Research Category FE	Yes	Yes	Yes	Yes	Yes	Yes
Team Size Control	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,147,106	1,437,097	1,437,097	1,147,106	1,437,097	1,437,097
Note:						*p<0.1; **p<0.05; ***p<0.01

Table 5: Stronger Effect Size with Greater Female Leadership

	Granular Disease Level Differences					
	Female Urogenital Diseases			Male Urogenital Diseases		
	(1)	(2)	(3)	(4)	(5)	(6)
Fem. Contributor	0.0005			−0.001**		
	(0.001)			(0.0005)		
Fem. First Author		0.004***			−0.002***	
		(0.0003)			(0.0002)	
Fem. Last Author			0.009***			−0.003***
			(0.0003)			(0.0002)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Research Category FE	Yes	Yes	Yes	Yes	Yes	Yes
Team Size Control	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,147,106	1,437,097	1,437,097	1,147,106	1,437,097	1,437,097
Note:						*p<0.1; **p<0.05; ***p<0.01

Table 6: Interaction Effect

	Measures of Gendered Research			
	Female MeSH	Male MeSH	Female Urogenital Diseases	Male Urogenital Diseases
	(1)	(2)	(3)	(4)
Fem First Author	0.019***	−0.003***	0.007***	−0.003***
	(0.001)	(0.001)	(0.0003)	(0.001)
Fem. Last Author	0.016***	−0.003***	0.007***	−0.003***
	(0.001)	(0.001)	(0.0004)	(0.001)
Fem. First*Last Author	0.012***	−0.008***	0.003***	−0.008***
	(0.001)	(0.001)	(0.001)	(0.001)
Year FE	Yes	Yes	Yes	Yes
Research Category FE	Yes	Yes	Yes	Yes
Team Size Control	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Observations	1,437,097	1,437,097	1,437,097	1,437,097
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 7: Female Leadership and Gender Equality

	Gendered MeSH	
	Female MeSH	Male Mesh
	(1)	(2)
Last Female Author	0.013***	−0.002
	(0.002)	(0.002)
High Gender Equality	0.016	−0.009
	(0.028)	(0.029)
Last Female Author*High Gender Equality	0.013***	−0.005**
	(0.002)	(0.002)
Year FE	Yes	Yes
Research Category FE	Yes	Yes
Team Size Control	Yes	Yes
Country FE	Yes	Yes
Observations	1,437,097	1,437,097
Note:		*p<0.1; **p<0.05; ***p<0.01

Discussion and Next Steps

I have demonstrated that female scientist are more likely to create science that focuses on female health - and all three conjectures stated in the hypothesis section stand up to the empirical evidence. I find an association between female scientist and female biomedicine; the association is stronger when a woman is the lead author on a paper; greater gender-equality is associated with female scientists working on diseases or conditions that disproportionately affect women. The more interesting idea cited in the background research - that increasing the number of female researchers will ameliorate gender bias in biomedicine - doesn't readily follow from the data.

Another tantalizing idea mentioned in the introduction - that women have knowledge from personal experience that allows them to innovate for other women - doesn't necessarily follow from my empirical results. I do not make a strong inference about why the effect of a women scientist is stronger when the woman is in a leadership role. It could be that promising female researchers are steered toward such work by their colleagues. It is possible that once women enter specific sub-discipline and beginning their work, men sort into other topics. Nonetheless, the presence of women seems to have contributed toward the growth of research into female-specific diseases. If we focusing strictly on lead authors, more women has meant more female focused biomedcine in the aggregate.

Although I am modest about what is actually driving this effect, there are some implication that follow directly from this finding. First, bias against women in STEM doesn't just impact individual careers (Lerchenmueller and Sorenson, 2018). It may actually impact the course of science and our understanding of women's health. Once it is established that the presence of female leadership in science is associated with female-focused biomedicine, it makes it hard to argue that these innovations could not have been eariler because of some constraint in the nature of biomedical science.

This result also adds to the research on the affects of diversity on group performance (Østergaard et al., 2011). This line of research studies many topics, but ignores the affect of group diversity on the nature of outputs produced (Østergaard et al., 2011). My result demonstrates that efforts to increase society wide gender equality may have second order effects that allow women scientist to express their preferences, which may be to innovate in female biomedicine. Since gender bias has been documented at all stages of science, it is worth exploring specifically how low gender equality countries manifest these biases.

More to the point, the vindication of hypothesis 3 suggests that the broader cultural milieu in which science is created impacts what (or how much) of a certain type of science is created. Again, it is interesting to divorce this from simply the rate of innovation, and try to see how demographics and societal wide

attitudes influences what actually gets invented. If gender diversity in biomedicine can impact the scope of innovations, should institutions only care about the rate of innovation? Is lower rate of innovation on a neglected topic actually better than fasterinnovation prefered on a topic that that is well-studied? There are more questions than answers, but this is could be a fruitful line of research.

This study adds to the existing literature in another important way. It provides some empirical evidence for those that wish to increase the numbers of women in STEM. In this respect, this result could be used to argue (with some caveats and qualifying statements) that increasing gender equality may have broader implications for female health outcomes than previously suspected. There is still lots to be explored with this data set. First and most obvious, it remains to be seen if this result will hold once more years are added. Second, it may useful to get a unique author ID and test if career age of the female researcher is associated with research on women (unique author ID exists in the Medline data base from 2015 onward). Such a unique author ID would allow us the follow early stage researcher to measure how they select into certain labs and/or research areas as their career progressed. Lastly, a promising test of sorting within researchers areas follows from the MeSH tree codes generated from the MTI output. For example, we can use the MeSH ontology to test for sorting across broad- and fine-grained research areas. Broader research areas being citations that return MeSH with level two depth within the ontological tree - e.g., vascular diseases [C14.907], and diabetes mellitus [C19.24] (see Appendix B for a visual). Whereas, fine-grained research would return tree codes such as atrial flutter [C14.280.067.248], and diabetes mellitus, Type 2, Lipoatrophic [C19.246.300.500]. If this sorting is gendered in any way, it may help explain why women are less likely to file biomedical patents. The assumption being that cutting edge research in science is now always very specialized. Furthermore, it may be the case that women are more interested in research on humans vs animals, and more experimental and cutting edge research is driven on animals models (the MeSH returns "Human" and "Animal" tags as well, allowing use to test for this sorting). I leave this as an open question.

In conclusion, I have demonstrated that female scientist are more likely to create science that focuses on female health. This is interesting not only for the policy implications, but it shows that it may be useful to think about how the nature of scientific output may depend on who the demographics of who is creating the science.

Appendix A - Robustness Check

Table 8: Robustness Check

	Gendered Keyword Match					
	Female Match			Male Match		
	(1)	(2)	(3)	(4)	(5)	(6)
Fem Contributor	0.011^{***}			-0.001^{*}		
	(0.001)			(0.001)		
Fem. First Author		0.026^{***}			-0.001^{***}	
		(0.001)			(0.0003)	
Fem. Last Author			0.024^{***}			-0.001^{***}
			(0.001)			(0.0003)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Research Category FE	Yes	Yes	Yes	Yes	Yes	Yes
Team Size Control	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,147,106	1,437,097	1,437,097	1,147,106	1,437,097	1,437,097

Note: *p<0.1; **p<0.05; ***p<0.01

Appendix B - MeSH Ontology

- Anatomy [A]
- Organisms [B]
- Diseases [C]
  - Infections [C01]
  - Neoplasms [C04]
  - Musculoskeletal Diseases [C05]
  - Digestive System Diseases [C06]
  - Stomatognathic Diseases [C07]
  - Respiratory Tract Diseases [C08]
  - Otorhinolaryngologic Diseases [C09]
  - Nervous System Diseases [C10]
  - Eye Diseases [C11]
  - Male Urogenital Diseases [C12]
  - Female Urogenital Diseases and Pregnancy Complications [C13]
  - Cardiovascular Diseases [C14]
  - Hemic and Lymphatic Diseases [C15]
  - Congenital, Hereditary, and Neonatal Diseases and Abnormalities [C16]
  - Skin and Connective Tissue Diseases [C17]
  - Nutritional and Metabolic Diseases [C18]
  - Endocrine System Diseases [C19]
  - Immune System Diseases [C20]
  - Disorders of Environmental Origin [C21]
  - Animal Diseases [C22]
  - Pathological Conditions, Signs and Symptoms [C23]
  - Occupational Diseases [C24]
  - Chemically-Induced Disorders [C25]
  - Wounds and Injuries [C26]
- Chemicals and Drugs [D]
- Analytical, Diagnostic and Therapeutic Techniques, and Equipment [E]
- Psychiatry and Psychology [F]
- Phenomena and Processes [G]
- Disciplines and Occupations [H]
- Anthropology, Education, Sociology, and Social Phenomena [I]
- Technology, Industry, and Agriculture [J]
- Humanities [K]
- Information Science [L]
- Named Groups [M]
- Health Care [N]

- Male Urogenital Diseases [C12]
  - Genital Diseases, Male [C12.294]
  - Pelvic Floor Disorders [C12.483]
  - Tuberculosis, Urogenital [C12.672]
  - Urogenital Abnormalities [C12.706]
  - Urogenital Neoplasms [C12.758]
    - Genital Neoplasms, Male [C12.758.409]
    - Urologic Neoplasms [C12.758.820]
  - Urologic Diseases [C12.777]
- Female Urogenital Diseases and Pregnancy Complications [C13]
  - Female Urogenital Diseases [C13.351]
    - Genital Diseases, Female [C13.351.500]
    - Pelvic Floor Disorders [C13.351.625]
    - Tuberculosis, Urogenital [C13.351.750]
    - Urogenital Abnormalities [C13.351.875]
    - Urogenital Neoplasms [C13.351.937]
      - Genital Neoplasms, Female [C13.351.937.418]
        - Fallopian Tube Neoplasms [C13.351.937.418.365]
        - Ovarian Neoplasms [C13.351.937.418.685]
        - Uterine Neoplasms [C13.351.937.418.875]
        - Vaginal Neoplasms [C13.351.937.418.937]
        - Vulvar Neoplasms [C13.351.937.418.968]
      - Urologic Neoplasms [C13.351.937.820]
    - Urologic Diseases [C13.351.968]
  - Pregnancy Complications [C13.703]
- Cardiovascular Diseases [C14]

Appendix C - Manually Indexed MeSH vs MTI Generated MeSH

Despite the fact that data on manually indexed MeSH is available from Medline, it cannot be used to captured gendered research. The main purpose of the manually indexed medical subject headings is to help researchers find related research. It is true that human indexers are instructed to review the title and abstracts of a paper and assigns tags for diseases that only effect one gender - i.e., "male" tag for research on prostate cancer and "female" tag for endometrioses - effectively giving us a measure of gendered research. But human indexers also review the body of the paper and assign "male" and "female" tags according to the gender of the subjects used in experiments (MEDLINE Indexing Online Training, 2019). This is done even if the paper is not studying a gendered disease. More to the point, if a paper studies the impact of statins on cardiovascular disease, and the subjects used are all male, a human indexer will apply a "male" check tag to the citation. One might think such a gender skew is an anachronism that doesn't apply more contemporary research. But this would be too optimistic. Even in the mid-aughts it was not uncommon to find only male subjects being used to study diseases (Kim et al., 2010). For this reasons, using the manually indexed MeSH is untenable - indeed it perpetuates the exact bias we are trying to circumvent. Below I provide a screen shot of a paper studying a non-gendered disease. Note that the human indexers attached a male tag to the citation because the experiment was done on males.

Clinico-pathological Correlations of Odontogenic Tumors: Some Critical Observations Based on a 20 Year Institutional Study and a Comprehensive Review of Literature

Shaheen Syed [1], Karla M Carvalho [1], Anita Spadigam [1], Anita Dhupar [1]

Affiliations + expand

PMID: 31745045 DOI: 10.4103/ijdr.IJDR_579_17

Free article

Abstract

Context: Odontogenic tumors (OTs) represent a rare subset of pathologies of the oral and maxillofacial region. The classification of OTs has undergone several changes over the years due to a lack of uniform international identification criteria. The histomorphological similarity and the many variations in behavioral patterns elaborated by these lesions warrant research.

Aims: Using the update from the fourth edition of the World Health Organisation Classification of Head and Neck Tumors (2017), this dental institution carried out an epidemiological study on OTs in the state of Goa (India) and compared the data obtained with similar studies on OTs done within India.

Figure 1: non-gendered diseases

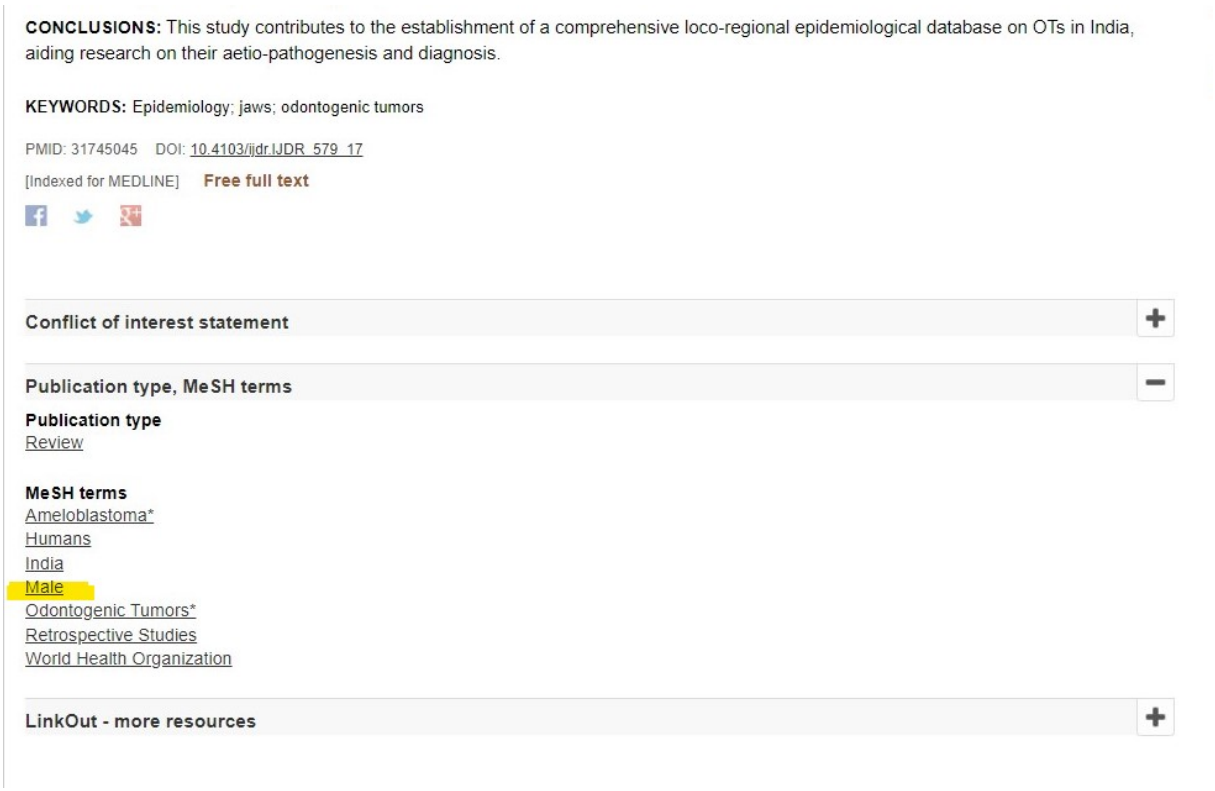

Figure 2: Manually index MeSH returns male tag. This would lead to measurment error when trying to capture gendered research

I take the title and abstract of the same citation and run it through the medical text indexer in interactive mode[5]. I highlighted the tree codes so you can see what the actual MTI output looks like. This may give some idea on how I approaced coding the C13/C12 urogenital diseases level differences.

Note that there is no male MeSH generated because this research is not gendered in the way I am trying to capture. We see that using the MTI gives a better measure because it avoids human indexers adding male/female tags based on the gender of the subjects used in the experiment.

Limitations of the Medical Text Indexer

My use of the MTI to generated MeSH is not above reproach, so here I acknowledge some limitations. The argument that the MTI's output gives a perfect measurement of gendered research rests on two assumptions, both of which must be true for me to get an unbiased measurement. First, it must be the case that any research on diseases that impact both men and women returns (1) no "male" or "female" MeSH, or (2) both "male" and "female" MeSH. If there is any information in the abstract or title that

[5]https://ii.nlm.nih.gov/Interactive/MTI/mti.shtml

Input Text:

```
Clinico-pathological correlations of odontogenic tumors: Some critical observations based on a 20 yea
Odontogenic tumors (OTs) represent a rare subset of pathologies of the oral and maxillofacial region.
variations in behavioral patterns elaborated by these lesions warrant research.
AIMS:
Using the update from the fourth edition of the World Health Organisation Classification of Head and
India.
```

Results:

```
Command: MTI -default_MTI -showTreecodes -justFacts
Version: 2018 Version of MeSH Used to Generate Recommendations

00000000|Humans|B01.050.150.900.649.313.988.400.112.400.400|C0086418|33538
00000000|World Health Organization|N03.540.514.718.800|C0043237|8892
00000000|Odontogenic Tumors|C04.557.695|C0028880|32538
00000000|Head and Neck Neoplasms|C04.588.443|C0018671|5145
00000000|Epidemiologic Studies|E05.318.760.500;N05.715.360.775.175;N06.850.520.450.500|C0002783|3846
00000000|Head|A01.456|C0018670|3135
00000000|India|Z01.252.245.393|C0021201|156
```

Figure 3: MTI text analysis does not return a male tag

only relates to one gender - but the research topic is on diseases that impacts both men and women - then the MTI will return a male or female tag, leading to a measurement error. The assumptions listed above don't hold for medical case studies. Consider a sample abstract of a medical case study below:

> "An 18-year-old white man was admitted to the Osler Medical Service with the chief complaint of back pain. Two weeks prior to admission, the patient developed diffuse and aching upper back pain. Over the next couple of days, he also developed severe anterior chest pain that was somewhat pleuritic in nature but diffuse and extending bilaterally into the shoulders. One week prior to admission, he developed intermittent fevers and night sweats. . . "

In this case, the MTI will return a male tag and no female tag, giving us a measurement error. Upon manual inspection of the data set, these case studies were the only types of research that produced erroneous male/female MeSH (at least for my purposes - it works fine if you consider the actual purpose of medical indexing: to help researchers find related research). I acknowledge this limitation and offer two robustness checks. First, I could process just the titles and leave the abstract out. This may get us a better measurement, but it is far from perfect. Consider the following title of a case study: "Headache, fever and back pain in a 16-year-old boy". Once again the MTI - and manual human indexing - will return

a male tag on a disease that doesn't only impacts males. Perhaps the only way to circumvent this problem would be to pre-process titles and abstracts in a way that does bias the MTI results on case studies.

Appendix D - Gendered Dictionary Keywords

Female Keywords

• Aborticide • Abortion • Abortus • Abruptio placentae • Adenomyosis • Amenorrhea • Amenorrheal • Amenorrheic • Amenorrhoeal • Amenorrhoeic • Anovulation • Antiestrogen • Areola • Areolae • Areolar • Areolas • Areolate • Areolation • Artificial Insemination • Bartholin's glands • Birth • Birth control • Birthed • Birthing • Births • Blastocyst • Blastosphere • Blastula • Blastulae • Blastular • Blastulas • Blastulation • Breast • Breast Cancer • Breast cyst • Breast reconstruction • Breast-feed • Breastfed • Breastfeed • Breastfeeding • Breasts • Breech • Brenner tumor • C section • C-section • Casesarian • Cervical • Cervical Cancer • Cervical canal • Cervical cerclage • Cervices • Cervix • Cervixes • Cesarean • Cesarean section • Child-bearing • Chorea gravidarum • Chorioamnionitis • Chorioamnionitis • Clitoral • Clitoral hood • Clitoric • Clitoridean • Clitoridectomies • Clitoridectomy • Clitoris • Colpitis • Colposcopy • Colpotomy • Culdoscopy • Diaphragm • Diaphragms • Dysmenorrhea • Dysmenorrheal • Echogenic bowel • Eclampsia • Eclamptic • Ectopic pregnancies • Ectopic pregnancy • Embryo • Embryonic • Embryos • Endocrine system • Endocrinology • Endometrial ablation • Endometrial hyperplasia • Endometrial neoplasms • Endometrioid carcinoma • Endometriosis • Endometritis • Endometrium • Estrogen • Estrus • Estrus cycle • Extrauterine pregnancy • FGM • Fallopian tube • Fallopian tube neoplasms • Fallopian tubes • Female Genital Mutilation • Female circumcision • Female condom • Female genital tuberculosis • Fetal • Fetal macrosomia • Feticide • Fetoscopy • Fetus • Fetuses • Fimbria • Fimbriae • Fimbrial • Foetus • G spot • G-spot • GYN • Graafian follicles • Grafenberg spot • Granulosa cell tumor • Grfenberg spot • Gynatresia • Gynecologic • Gynecological • Gynecologist • Gynecologists • Gynecology • HELLP syndrome • Hematocolpos • Hematometra • Hereditary breast and ovarian cancer syndrome • Hot flash • Hot flashes • Hot flush • Hydrocolpos • Hydrops fetalis • Hymen • Hymenal • Hyperemesis gravidarum • Hysterectomies • Hysterectomy • Hysteroscopy • Hysterotomy • IUD • Impregnation • Infibulation • Intrauterine device • Labia • Labia majora • Labia minora • Labium • Lactate • Lactating • Lactation • Leukorrhea • Luteoma • Maternal • Mammaplasty • Mammary • Mammary Gland Lobules • Mammary gland • Mammectomy • Mammogram • Mammography • Mastectomies • Mastectomy • Meigs syndrome • Menopause • Menorrhagia • Menses • Menstrual • Menstrual blood • Menstrual cycle • Menstruating • Menstruation • Metrorrhagia • Miscarriage • Mons pubis • Montes pubis • Mother • Mothers • Multibirth • Multiovulate • Multiovulated • Myometrium • Nuchal cord • Oestrus • Oligohydramnios • Oocyte • Oocytes •

Oogonium • Oophoritis • Oosphere • Ova • Ovarian • Ovarian Cancer • Ovarian cyst • Ovarian cysts • Ovarian hyperstimulation sndrome • Ovarian neoplasms • Ovariectomy • Ovaries • Ovary • Oviducal • Oviduct • Oviductal • Oviducts • Ovulate • Ovulated • Ovulating • Ovulation • Ovulatory • Ovum • PCOS • Pap smear • Pap test • Papanicolaou test • Parametritis • Parovarian cyst • Pelvic inflammatory disease • Polycystic Ovary Syndrome • Post-Cesarean • Postabortion • Postpartum hemorrhage • Postpregnancy • Preeclampsia • Preeclamptic • Pregnancies • Pregnancy • Pregnant • Premature menopause • Prenatal • Prenatally • Preovulatory • Preterm birth • Primary ovarian insufficiency • Prophylactic mastectomy • Pseudovaries • Pseudovary • Pudenda • Pudendum • Puerperal infection • Pyelectasis • Pyometra • Radical mastectomy • Rectovaginal fistula • Salpingectomy • Salpingitis • Salpingo-oophorectomy • Salpingostomy • Segmental mastectomy • Skene's glands • Spontaneous abortion • Squamous intraepithelial lesions of the cervix • Stillbirth • Subcutaneous mastectomy • Symphysiotomy • Thecoma • Trachelectomy • Trophoblastic neoplasms • Turner syndrome • Uterine Fibroids • Uterine cancer • Uterine cervical dysplasia • Uterine cervical erosion • Uterine cervical incompetence • Uterine cervical neoplasms • Uterine cervicitis • Uterine inversion • Uterine myomectomy • Uterine prolapse • Uterine retroversion • Uterine rupture • Uterine sinus • Uterus • Uteruses • Vacuum curettage • Vagina • Vaginae • Vaginal cancer • Vaginal smears • Vaginas • Vaginismus • Vaginitis • Vasa previa • Vesicovaginal fistula • Vestibular bulb • Vestibular bulbs • Virilism • Vulva • Vulvae • Vulval • Vulvar • Vulvar cancer • Vulvar lichen sclerosus • Vulvar neoplasms • Vulvar vestibulitis • Vulvas • Vulvate • Vulvectomy • Vulviform • Vulvitis • Vulvodynia • Vulvovaginal candidiasis • Vulvovaginitis • Wet-nurse • Wet-nurses • Womb

Male Keywords

• Alport syndrome • Androgenetic alopecia • Aspermia • Asthenozoospermia • Azoospermia • Azoospermic • BPH • Balanitis • Balanoposthitis • Bald • Baldness • Benign prostatic hypertrophy • Bulbourethral glands • Castration-resistant prostatic neoplasms • Circumcised • Circumcision • Circumcisions • Corpus cavernosum • Cowper's glands • Cremaster muscle • Cryptorchid • Cryptorchidism • Cryptorchidism • Ejaculatory duct • Enlarged prostate • Epididymal • Epididymides • Epididymis • Epididymitis • Erectile dysfunction • Erectile tissue • Foreskin • Fournier gangrene • Glans penis • Gonadal dysgenesis • Gonadoblastoma • Haemophilia • Hematocele • Hematospermia • Hemospermia • Hydrocele • Hypospadiac • Hypospadias • Impotence • Impotent • Infecund • Infecundity • Inseminate • Insemination • Interseminal • Klinefelter syndrome • Low sperm count • Male genital tuberculosis • Male pattern baldness • Micropenis • Microphallus • Oligospermia • Orchiectomy • Orchiopexy • Orchitis • Paraphimoses • Paraphimosis • Paraphimotic • Paternal • Penes • Penial • Penile • Penile cancer • Penile induration • Penile neoplasms • Penis • Penises • Peyronie's disease • Phimoses • Phimosis • Phimotic • Premature ejaculation • Prepuce • Preseminal • Priapism • Priapismic • Prostate • Prostate cancer • Prostate gland • Prostatectomy • Prostatic hyperplasia • Prostatic neoplasms • Prostatic sinus • Prostatic sinuses • Prostatitic • Prostatitis • Puboprostatic • Puboprostatic ligament • Scrota • Scrotal • Scrotum • Scrotums • Semen • Seminal • Seminal gland • Seminal glands • Seminal vesicle • Seminal vesicles • Seminality • Seminally • Seminoma • Seminomas • Seminomata • Sertoli cell-only syndrome • Sperm motility • Spermatic cord torsion • Spermatogenesis • Spermatogenetic • Spermatozoa • Spermicidal • Spermicide • Squamous cell cancer • Teratozoospermia • Testes • Testicles • Testicular cancer • Testicular cord • Testicular neoplasms • Testicular torsion • Testis • Testis tumor • Testosterone • Transrectal ultrasound • Varicoceles • Variococele • Vas deferens • Vasa deferentia • Vasculogenic impotence • Vasectomy • Vasovasostomy • Y chromosome • Y chromosomes

References

Aronson, A. R., Mork, J. G., Gay, C. W., Humphrey, S. M., and Rogers, W. J. 2004. The NLM Indexing Initiative's Medical Text Indexer, *Studies in health technology and informatics*, vol. 107, no. Pt, 268–72

Azoulay, P., Ganguli, I., and Graff Zivin, J. 2017. The mobility of elite life scientists: Professional and personal determinants, *Research Policy*, vol. 46, no. 3, 573–90

Barlow, R. 2018. Why Medical Research Often Ignores Women | BU Today, *Boston University*

Bell, A., Chetty, R., Jaravel, X., Petkova, N., and Van Reenen, J. 2019. Who Becomes an Inventor in America? The Importance of Exposure to Innovation*, *The Quarterly Journal of Economics*, vol. 134, no. 2, 647–713

Holman, L., Stuart-Fox, D., and Hauser, C. E. 2018. The gender gap in science: How long until women are equally represented?, (C. Sugimoto, Ed.), *PLOS Biology*, vol. 16, no. 4, e2004956

Jensen, K., Kovács, B., and Sorenson, O. 2018. Gender differences in obtaining and maintaining patent rights, *Nature Biotechnology*, vol. 36, no. 4, 307–9

Kahn, S. and Ginther, D. 2017. Women and Stem, *NBER Working Paper*, vol. No. w23525, no. No. w23525

Kim, A. M., Tingen, C. M., and Woodruff, T. K. 2010. Sex bias in trials and treatment must end: Gender inequalities in biomedical research are undermining patient care. In the first of three related pieces, Alison M. Kim, Candace M. Tingen and Teresa K. Woodruff call on journals, funding agencies and researchers to give women parity with men, in studies and in the clinic, *Nature*, vol. 465, no. 7299, 688+

Klein, J. 2019. Fighting the Gender Stereotypes That Warp Biomedical Research - The New York Times, *New York Times*, Advance Access published May 2019

Koning, R., Samila, S., and Ferguson, J.-P. 2019. Female Inventors and Inventions: Social Science Research Network ID 3401889

Konrad, A. M., Kramer, V., and Erkut, S. 2008. Critical Mass: The Impact of Three or More Women on Corporate Boards, *Organizational Dynamics*, vol. 37, no. 2, 145–64

Lerchenmueller, M. J. and Sorenson, O. 2018. The gender gap in early career transitions in the life

sciences, *Research Policy*, vol. 47, no. 6, 1007–17

Little, B. 2014. Delivering "The Pill" Wasn't Easy, *National Geographic News*

Long, J. S. and Fox, M. F. 1995. Scientific Careers: Universalism and Particularism, *Annual Review of Sociology*, vol. 21, no. 1, 45–71

Lukong, K. E. 2017. Understanding breast cancer - The long and winding road, *BBA clinical*, vol. 7, 64–77

MEDLINE Indexing Online Training. 2019. Advance Access published December 2019

Mork, J. G., Jimeno-Yepes, A., and Aronson, A. R. 2013. The NLM Medical Text Indexer System for Indexing Biomedical Literature, in *BioASQ@CLEF*

Murray, F. and Graham, L. 2007. Buying science and selling science: Gender differences in the market for commercial science, *Industrial and Corporate Change*, vol. 16, no. 4, 657–89

Nielsen, M. W., Alegria, S., Börjeson, L., Etzkowitz, H., Falk-Krzesinski, H. J., Joshi, A., Leahey, E., Smith-Doerr, L., Woolley, A. W., and Schiebinger, L. 2017a. Opinion: Gender diversity leads to better science, *Proceedings of the National Academy of Sciences*, vol. 114, no. 8, 1740–2

Nielsen, M. W., Andersen, J. P., Schiebinger, L., and Schneider, J. W. 2017b. One and a half million medical papers reveal a link between author gender and attention to gender and sex analysis, *Nature Human Behaviour*, vol. 1, no. 11, 791–96

Office of Research on Women's Health. 2016. Report of the advisory committee on research on women's health: National Institute of Health

Programme, U. N. D. 2016. Human Development Index (HDI):

Srivastava, S. B. and Sherman, E. L. 2015. Agents of Change or Cogs in the Machine? Reexamining the Influence of Female Managers on the Gender Wage Gap, *American Journal of Sociology*, vol. 120, no. 6, 1778–1808

Torchia, M., Calabrò, A., and Huse, M. 2011. Women Directors on Corporate Boards: From Tokenism to Critical Mass, *Journal of Business Ethics*, vol. 102, no. 2, 299–317