

1 **Factors affecting sex-related reporting: a cross-disciplinary bibliometric**  
2 **analysis of medical research**

3 Cassidy R. Sugimoto PhD<sup>1</sup>, Yong-Yeol Ahn PhD<sup>1</sup>, Elise Smith PhD<sup>2</sup>, Benoit Macaluso MLIS<sup>3</sup>,  
4 Vincent Larivière PhD<sup>2,3\*</sup>

5 <sup>1</sup>School of Informatics, Computing, and Engineering, Indiana University Bloomington, USA.

6 <sup>2</sup>École de bibliothéconomie et des sciences de l'information, Université de Montréal, Canada.

7 <sup>3</sup>Observatoire des sciences et des technologies, Centre interuniversitaire de recherche sur la  
8 science et la technologie, Université du Québec à Montréal, Canada

9 \*Correspondence to: vincent.lariviere@umontreal.ca

10

11 **Keywords:** sex-related reporting, gender disparities, science policy, health research

12 Running title: Sex-related reporting in medical research

13 Word count: 4160

14 **Funding:** This study was funded by the Canada Research Chairs program (grant #950-231768).  
15 CRS performed this work while supported under the Independent Research & Development  
16 program at the National Science Foundation.

17 **Author contributions:** Conceived and designed the experiments: CRS, ES, VL. Performed the  
18 experiments: CRS, YYA, BM, VL. Analyzed the data: CRS, YYA, BM, VL. Wrote the paper:  
19 CRS, YYA, ES, VL.

20

21

## 22 **ABSTRACT**

### 23 **Background**

24 Clinical and pre-clinical studies have shown that there are sex-based differences at the genetic,  
25 cellular, biochemical, and physiological levels. Despite this, numerous studies have  
26 demonstrated a lack of inclusion of female populations into medical research. These disparities  
27 in sex inclusion are further problematized by the lack of sex reporting: that is, describing the  
28 population under study. Disparities in inclusion of both sexes in medical research significantly  
29 reduces the utility of the results of medical research for the entire population. The lack of sex  
30 reporting can be problematic for the translation of research from the pre-clinical to clinical and  
31 applied health settings. Large-scale studies are needed to identify the degree of sex-related  
32 reporting and where disparities are more prevalent. Furthermore, there are several studies  
33 showing the dearth of female researchers in science, yet few have evaluated whether a lack of  
34 women in science may be related to disparities in sex inclusion and reporting.

### 35 36 **Methods**

37 This paper analyses sex-related reporting in medical research, based on a set of more than 11.5  
38 million papers indexed in Web of Science and PubMed between 1980 and 2016 and using sex-  
39 related Medical Subject Headings as a proxy for sex reporting. Descriptive statistics and  
40 regression analyses are used to analyze these data.

### 41 42 **Results**

43 Despite an increase in sex-related reporting between 1980 and 2016 in clinical medicine (59% to  
44 67%) and public health research (36% to 69%), sex remains largely underreported in biomedical  
45 research (31% in 2016). Furthermore, papers with female first and last authors have a higher  
46 probability of reporting sex, with an odds ratio of 1.26 (95% CI: 1.24-1.27) and sex-related  
47 reporting is associated with publications in journals with low impact factors. For instance, for the  
48 publications in 2016, sex-related reporting of both male and female is associated with the  
49 reduction of -0.51 (95% CI: -0.54, -0.47) in impact factors.

### 50 51 **Interpretation**

52 This paper suggests that the current gender disparities in the scientific workforce and lack of  
53 policies on sex-reporting at the journal and institutional level may inhibit effective research  
54 translation from bench to clinical studies.

## 55 56 **RESEARCH IN CONTEXT**

### 57 58 **Evidence before this study**

59 Literature review searches were conducted in June 2016 (and periodically thereafter) on several  
60 bibliometric databases, including PubMed, Web of Science, and Google Scholar, using terms  
61 such as “sex reporting”, “sex analysis”, “sex inclusion” as well as terms on “gender bias”,  
62 “gender disparities”, and “sex factors”. The latter terms were particularly analyzed in reference  
63 to bibliometric terms (e.g., “citation” and “author”). The queries revealed several hundred  
64 articles on related topics, primarily reinforcing sex-based differences in medicine and the  
65 underrepresentation of women in science. These studies demonstrated that there are strong sex-  
66 based differences at the genetic, cellular, biochemical, and physiological levels and argue for the

67 construction of policies for greater sex-related reporting and analysis in medical research. Sex-  
68 related reporting has been shown to be low, but increasing. However, extant studies are often  
69 monodisciplinary (or cover only a few specific specialties or diseases) and fail to account for the  
70 translation from biomedical, to clinical, to public health research. This has potentially negative  
71 effects as research done on one sex in the biomedical phase are then translated and used on  
72 patients of the opposite sex in public health research. Furthermore, there is a growing body of  
73 research suggesting a relationship between the gender of the researchers and the outcomes of the  
74 research.

75

#### 76 **Added value of this study**

77 Our findings provide clear evidence of the growth of sex-related reporting in research and how it  
78 varies across medical disciplines and specialties. Clinical specialties report on sex much more  
79 than biomedical specialties, with fertility, obstetrics and gynecology, and urology having the  
80 highest incidence of sex-related reporting, and hematology, immunology, and pharmacy having  
81 the lowest. Controlling for confounding factors, female first or last authors have a higher  
82 probability of sex-related reporting, and are more likely to report studying females or both sexes,  
83 and journals with the high impact factors are less likely to report sex. This provides a  
84 contemporary and comprehensive analysis that complements earlier studies of rates of sex-  
85 related reporting and provides a novel extension of research demonstrating the relationship  
86 between sex-related reporting and author gender.

87

#### 88 **Implications of all the available evidence**

89 There has been a strong increase in sex-related reporting, particularly in clinical research and  
90 public health, but sex remains widely underreported in biomedical studies. This can be addressed  
91 through policies at several levels: funding agencies should mandate sex-related reporting in  
92 proposals and journal editors should insist upon sex-related reporting in submissions. Sex-related  
93 reporting should be a necessary requirement for ethical and replicable medical science.  
94 Furthermore, this research suggests several consequences of the demographic composition of the  
95 scientific workforce and the distribution of labor on scientific teams. Women are  
96 underrepresented in leadership positions and more likely to conduct experimentation than to be  
97 responsible for research design. Our research suggests that this is likely to be related to lower  
98 rates of sex-related reporting and analysis, particularly for female populations. Diversification of  
99 the scientific workforce is essential to produce the most rigorous and effective medical research.

100

## 101 **INTRODUCTION**

102 Sex matters in science. Numerous clinical and pre-clinical research studies have shown that there  
103 are sex-based differences at the genetic, cellular, biochemical, and physiological levels. Indeed,  
104 sex is at the source of numerous cellular variabilities, including rate of tissue re-generation (1),  
105 plaque formation (with critical implications for coronary artery disease) (2), and even  
106 susceptibility to neuronal cell starvation (3). Research on animal and human subjects has shown  
107 sexual dimorphism in cardiovascular disease, pulmonary issues, kidney problems, autoimmune  
108 disease, and various neurological conditions (4-5). Despite this, females have often been under-  
109 represented or excluded from research, with grave consequences. For example, the inadequate  
110 consideration of sex differences in pharmacokinetics and pharmacodynamics (6-7) has led to

111 disastrous results: of drugs withdrawn from the market from 1997 to 2001, 80% posed greater  
112 health risks for women than for men (8).

113  
114 A bias for male samples in pre-clinical research has often been justified by an alleged  
115 inconsistency caused by female oestrous cycles; the underlying rationale for this exclusion was  
116 that a homogeneous sample that limited diversity as much as possible would enable the isolation  
117 of key variables and lead to more coherent results. However, recent empirical research has  
118 shattered the myth of female variability, finding that males exhibit greater variability than  
119 females on a number of traits (9-10, 13-15).

120  
121 Recognizing that the costs of omission are far greater than any downside of inclusion, the 1993  
122 Revitalization Act mandated the increased enrollment of women in clinical trials for  
123 government-funded research. By 2013 more than half of all participants in National Institutes of  
124 Health (NIH)-funded clinical research studies were female (9) and there was a strong increase in  
125 sex-inclusive research. However, male bias during that same time increased in animal studies  
126 (10) and dominated research of cultured cells (11-12).

127  
128 The continued avoidance of sex-related reporting and analysis in pre-clinical studies reduces the  
129 ability to replicate research, gain knowledge on sexual dimorphism, and identify heterogeneity  
130 within female samples. Consequently, it also reduces effectiveness of research translation—  
131 potentially augmenting the risks—of clinical studies on human subjects. To address this, the NIH  
132 issued a policy in 2014 that called for balanced use of male and female cells and animals in  
133 preclinical studies, unless sex-specific exclusion could be rigorously justified (16).

134  
135 The sex of the research subject or sample is not the only place where sex matters in scientific  
136 research. Studies increasingly emphasize the importance of the demographic characteristics of  
137 the scientist and the interaction between scientists and those studied (35). For example, one study  
138 found that male laboratory technicians increased the stress of rodents under study, particularly  
139 female rodents (17). Furthermore, there is increasing evidence that the presence of female  
140 investigators may lead to increased sex analysis in research (18; 39).

141  
142 However, the extant literature fails to provide a contemporary and cross-disciplinary analysis of  
143 the degree of sex-related reporting across the health sciences—from biomedical, to clinical, and  
144 public health research—and the role of author gender in sex-related reporting. The present study  
145 seeks to address this gap.

## 146 147 **METHODS**

148 We contribute to this line of research with a large-scale analysis of 11.5 million articles. The  
149 goals of the paper are 1) to provide a comprehensive analysis of sex-related reporting across all  
150 specialties of biomedical, clinical, and public health research over the last 37 years; 2) to test the  
151 relationship between author gender and sex-related reporting in medical research; and 3) to  
152 examine factors that are associated to sex-related reporting in medical research. There is  
153 considerable ambiguity in the use of terms to describe sex-related reporting. Sex inclusion is  
154 often used to describe the inclusion of male and female populations in study and sometimes to  
155 refer exclusively to the inclusion of minority populations in a domain. Sex analysis is used to  
156 refer to the use of sex as an analytic variable in a study (thereby requiring the inclusion of both

157 sexes). Sex reporting is often used to denote the identification of the sex of the included  
158 population. In the present study, MeSH headings are used as a proxy for sex reporting. We  
159 therefore use the term “sex-related reporting” to denote studies that include the specified MeSH  
160 headings.

## 161 162 **PubMed**

163 Data from PubMed were downloaded from the U.S. National Library of Medicine bulk  
164 download website<sup>1</sup>. Raw XML data were transformed into a relational SQL database that allows  
165 for the compilation of bibliometric indicators. All Medical Subject Headings (MeSH) associated  
166 with sex (major and non-major topics) were used to retrieve papers that report sex (Table S1). In  
167 order to have mutually-exclusive categories of papers, we have categorized papers by reporting  
168 1) only female, 2) only male, 3) both sexes, or 4) no sex. Given the concerns that have been  
169 raised regarding the use of classification systems for examining sex in clinical and public health  
170 data (36-37), we conducted a validation exercise to check for false negatives and false positives  
171 in our data. Our analysis is based on the assumption that those studies that report on the sex of  
172 humans, animals, and cell cultures include an indicative sex-related MeSH. To test the use of  
173 MeSH for sex-related reporting, we used a specialties-based stratified sampling of articles and  
174 did and *not* include a sex-related MeSH (See Appendix). This has shown that, whereas MeSH  
175 serve as indicative of sex-related reporting, they cannot be used an indicator of sex analysis.

## 176 177 **Web of Science**

178 To obtain citation data, journal disciplinary classification, the Journal Impact Factor (JIF), and  
179 assign gender to authors, we matched papers indexed in PubMed with their equivalent record in  
180 Clarivate Analytics’ Web of Science (WoS). The matching of PubMed records with those of  
181 WoS was primarily conducted using three sets of matching keys: 1) Digital Object Identifiers  
182 (DOIs); 2) title, publication year, first author, and starting page; and 3) volume, publication year,  
183 first author, and starting page. Additional matching was also performed using the title,  
184 publication year, first page, and journal name, using a conversion table for journal names—based  
185 on the set of papers matched using the three abovementioned keys—as well as fuzzy logic was  
186 applied when titles were not identical. Over the 1980-2016 period, 88.2% (16,192,312 papers) of  
187 PubMed papers published in journals indexed by the WoS (N=18,349,143) were matched; this  
188 percentage increases from 81.9% in 1980 to 89.0% in 2016, mostly due to the greater presence of  
189 DOIs. Papers matched with WoS were attributed to a discipline and a speciality based on the  
190 classification developed for and used by the U.S. National Science Foundation. In total,  
191 11,572,428 papers were matched between PubMed and WoS over the 1980-2016 period, once  
192 limited to the field of Biomedical Research, Clinical Medicine, and Public Health (as per the  
193 National Science Foundation field and subfield classification) and to research and review  
194 articles. Public Health covers a majority of papers public health and health policy, as well as  
195 geriatrics and nursing, among others. Contrary to the WoS Subject Categories, this classification  
196 scheme classifies each journal into one discipline and one specialty. JIFs were corrected for the  
197 asymmetry between numerator and denominator (41), which means that only citations received  
198 by articles and reviews are counted in the numerator.

## 199 200 **Gender assignment**

<sup>1</sup> [https://www.nlm.nih.gov/databases/download/pubmed\\_medline.html](https://www.nlm.nih.gov/databases/download/pubmed_medline.html)

201 The WoS began indexing given names of researchers in 2008, which allows for the assignation  
 202 of a perceived gender to authors. Thus, for papers published between 2008 and 2016 and which  
 203 could be matched with PubMed ( $N=3,298,951$ ) we assigned gender of first and last authors—  
 204 which can be considered in medicine as dominant authorship positions (29)—using their names  
 205 following the assignment algorithm described in (26). More details on the algorithm, which has  
 206 also been used in (42-43), can be found in the supplementary materials and files of (26). The  
 207 algorithm assigned a gender to 72.4% of first authorships ( $N=2,387,311$ ) and 76.0% of last  
 208 authorships ( $N=2,508,420$ ). Names that remained unassigned were mostly due to initials (i.e.,  
 209 11.8% of first and 12.4% of last authors), with a small number of names that could not be  
 210 confidently assigned a gender (15.8% of first and 11.6% of last authors).

211  
 212 A brief note on terminology is warranted here. We use the term sex to discuss the samples or  
 213 populations under study, while we use gender to refer to the author on papers. Gender of authors  
 214 is determined by names, which provide—within a reasonable margin of error—the perceived  
 215 gender of the authors. This distinction is deceptively simple: the concept of ‘sex’ is usually  
 216 understood as involving biological attributes such as reproductive, hormonal, genetic, and  
 217 metabolic differentiation between male and female (30); gender, by contrast, is a concept that  
 218 includes cultural and psychosocial factors linked to sex but often determined as a type of  
 219 “embodied social structure” (31). However, because it is often difficult to assess what is due to  
 220 sex and what is due to gender or both, the notions are often conflated in medical research. For  
 221 example, there is a sex-based difference between a female’s auto-immune response which is  
 222 generally higher than that of males due to hormonal differences (32), but gender differentiation  
 223 may also modulate immune disorder because of external exposure (e.g., chemical, viral,  
 224 bacterial) (33). In this research, we will use the notion of sex to characterize populations,  
 225 samples, cells, etc. knowing that this may be linked to gender; conversely, we will use gender  
 226 when considering the authors of the research acknowledging that this is also related to sex.

227  
 228 **Regression analysis**

229 Starting from 3,298,951 papers, we first removed papers for which we could not determine the  
 230 gender of either first or last author ( $N = 1,192,430$ ). We then created two tables for single-author  
 231 papers ( $N = 87,824$ ) and multi-author ones ( $N = 2,018,697$ ) (Table 1). We used logistic  
 232 regression and OLS linear regression models to analyze the data. A full description of these  
 233 analyses can be found in the Appendix.

234  
 235 **Table 1.** Descriptive statistics of the sex-related reporting and the gender of the authors.

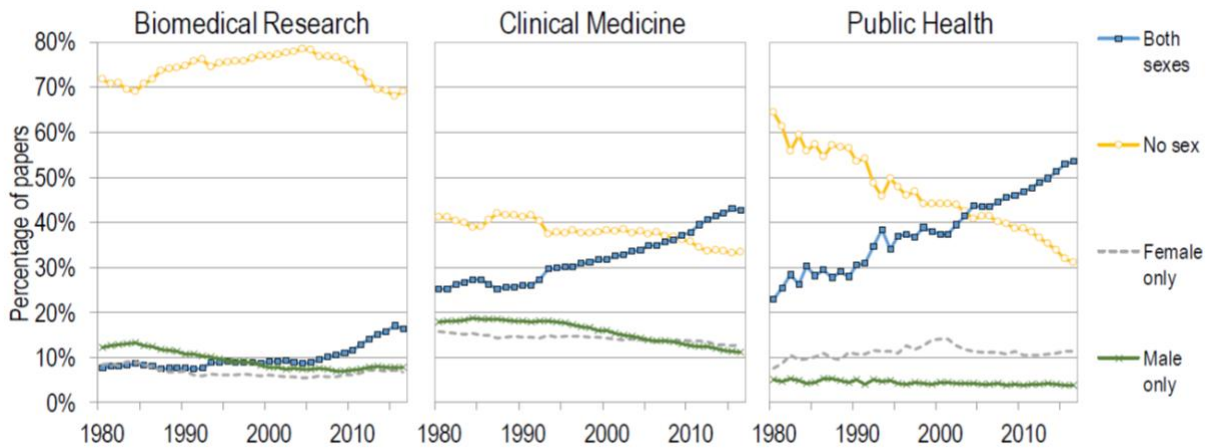
Sex reported?	Gender of the first and last authors				
	Total sample	FF	FM	MF	MM
Yes	1,127,989	180,136	305,738	147,174	494,941
No	890,708	117,959	246,861	119,971	405,917

236  
 237 The funders had no role in study design, data collection and analysis, decision to publish, or  
 238 preparation of the manuscript. CRS, YYA, BM, and VL had access to the data. All authors were  
 239 responsible for the decision to submit the manuscript.

241 **RESULTS**

242

243 There has been a dramatic increase in sex-related reporting in clinical medicine and public health  
 244 research (Figure 1). In 1980, only 36% of public health research reported on the sex of the  
 245 participants. By the late 1990s, the majority of studies reported on sex (69%) and a growing  
 246 number focused on female-only populations (from 8% to 11%). By 2016, the majority (54%) of  
 247 public health studies reported both male and female populations. In public health, single sex  
 248 studies focus more often on females than on males (11% vs. 4%). Clinical studies show an  
 249 increase of sex-related reporting from 59% to 67%—although until recently males were included  
 250 more often than female. The move to report both sexes occurred much later in clinical studies  
 251 than in public health: while more than half of papers in public health indicate sex-related  
 252 reporting in 2016, this percentage is at 43% in clinical medicine. Despite calls for reform, sex  
 253 remains underreported in biomedical research; the great majority of papers (nearly 70% in 2016)  
 254 fail to report on the sex of samples. However, in recent years, there has been a moderate increase  
 255 in the number of studies that incorporate both sexes, though this appears to be due to a decrease  
 256 in the number of single-sex studies, rather than an increase in any type of reporting.



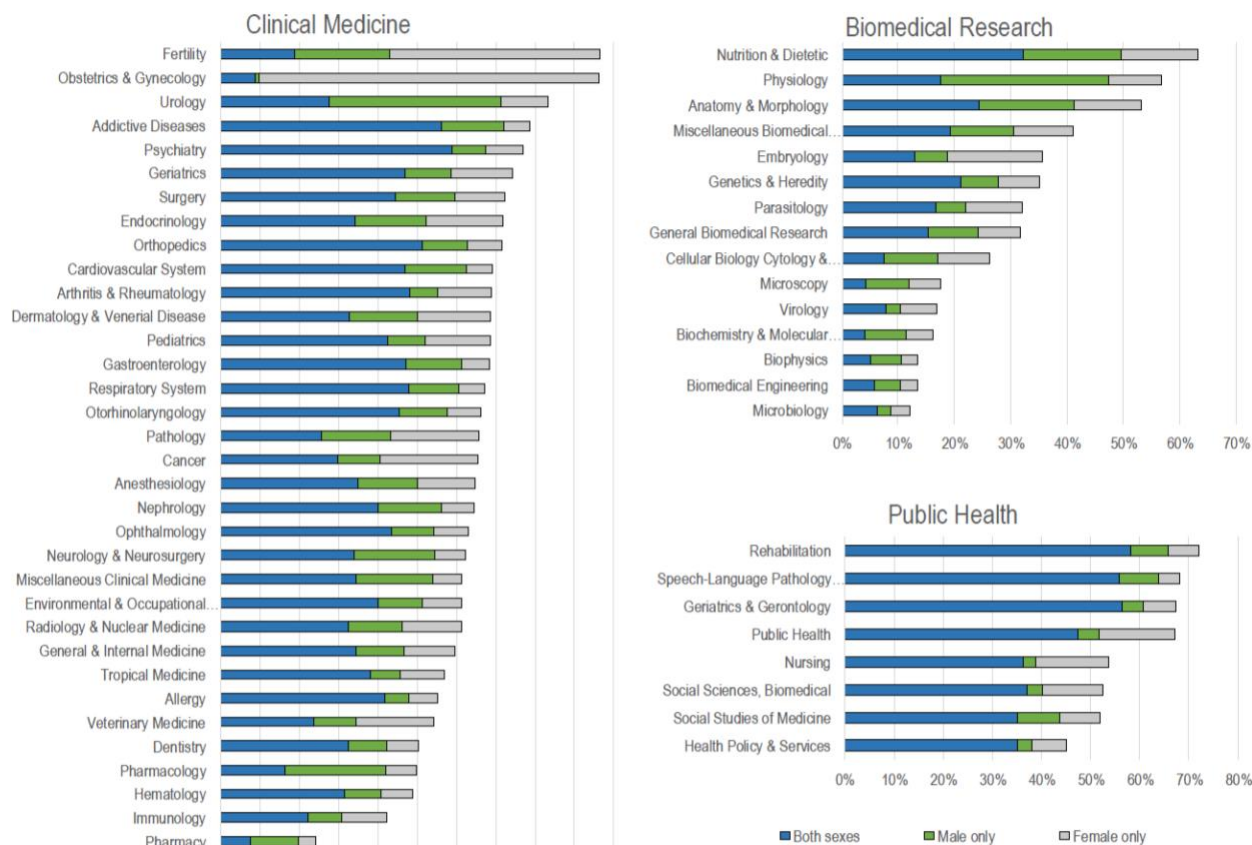
257

258

259 **Fig. 1.** Percentage of papers addressing sex (MeSH terms), by discipline, 1980-2016

260

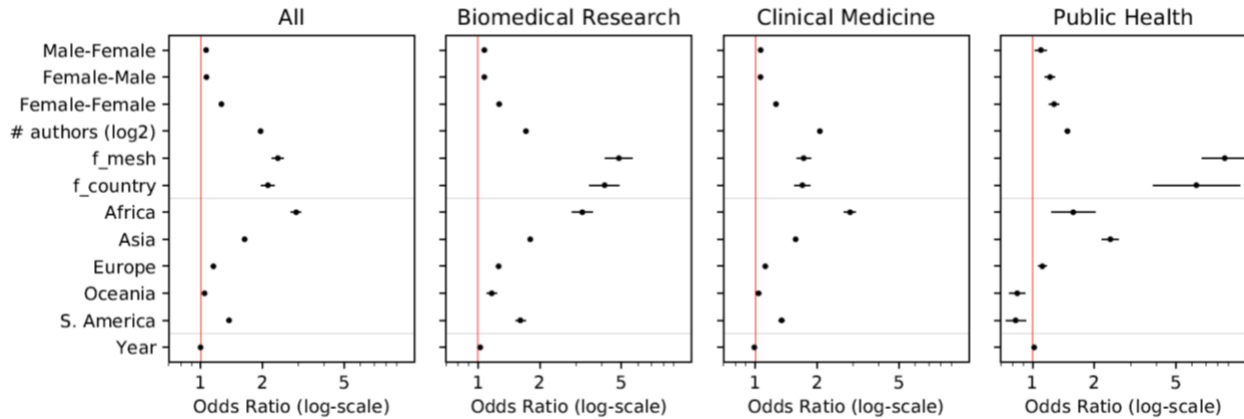
261 Fields are not equal when it comes to sex-related reporting (Figure 2). Fertility (97%), obstetrics  
 262 and gynecology (96%), and urology (83%) are among those disciplines with the greatest  
 263 incidence of sex-related reporting. Clinical medicine fields with a cellular or biochemical focus,  
 264 such as hematology (49%), immunology (42%), and pharmacy (24%), have the lowest levels of  
 265 sex-related reporting. This aligns with the distribution of sex-related reporting in the domain of  
 266 biomedical research, where only nutrition (63%), physiology (57%), and anatomy (53%) have a  
 267 majority of papers reporting on the sex of the population. Furthermore, in these disciplines,  
 268 males are studied more often than females. Public health research has the largest percentage of  
 269 sex-related reporting with a norm towards including both sexes in the analysis—54% in 2016.



270  
 271 **Fig. 2.** Percentage of papers addressing sex (MeSH terms), by specialty, 1980-2016  
 272

273 We estimated logistic regression models to study relationships between the gender of the authors  
 274 and sex-related reporting. The dependent variable of our models is the reporting (SR=1) or non-  
 275 reporting (SR=0) of sex (see Methods and Supplementary Materials for more details and  
 276 alternative models). The odds ratios of the key independent variables are shown in Figure 3.  
 277 Upon controlling the number of authors, representation of women in specific diseases (f\_mesh)  
 278 and in countries (f\_country), continents, year, and specialty areas, having female first or last  
 279 authors is positively associated with sex-related reporting. The effect size is the largest when  
 280 both first and last authors are female, with odds ratio of 1.26 (95% CI: 1.24-1.27). The number of  
 281 authors is also associated with the reporting of sex. Having twice as many authors corresponds to  
 282 the odds ratio of 1.96 (95% CI 1.94-1.97). There are also regional variations: compared with  
 283 North America, papers from all other regions, particularly Africa, are more likely to report sex.  
 284 This variation may stem from the different prevalence of research topics across regions rather  
 285 than biases or norms. Finally, the effect size of the ‘year’ variable is almost zero, suggesting that  
 286 most of the temporal variation may be explained by other factors, such as increasing number of  
 287 female authors and papers from outside the U.S.



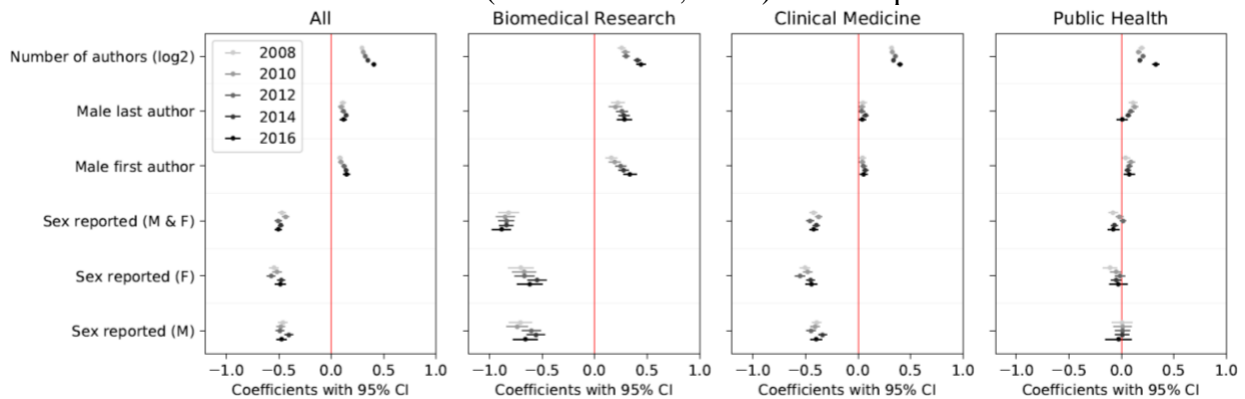


288  
289

290 **Fig 3.** Odds ratio of sex-related reporting from the logistic regression analysis. Throughout our  
291 models, the reference variable for the author’s gender combination is Male-Male and that for the  
292 geography is North America. The error bars represent 95% confidence intervals, which are  
293 smaller than the symbol in many cases. The leftmost plot shows the result from the aggregated  
294 dataset that include all three major disciplines (still controlling for all sub-disciplines), while the  
295 three following plots show results based on each major discipline separately. The effect of  
296 having female author(s) is positive across all cases. See Table S3-S10 for the regression tables,  
297 including those for the SR\_M, SR\_F, and SR\_B models.

298

299 Current incentive structures value placement of research in journal with high journal impact  
300 factors. However, high impact journals are not examples of best practices regarding sex-based  
301 reporting. Papers with sex-based reporting are more likely to appear in lower-impact journals  
302 than those without sex-based reporting, even when controlling for speciality of publication (Fig.  
303 4). For instance, for the publications in 2016, sex-related reporting of both male and female is  
304 associated with the reduction of -0.51 (95% CI: -0.54, -0.47) in the impact factor.



305  
306

307 **Fig. 4.** The effect sizes of independent variables on the impact factor of the journal. The error  
308 bars represent 95% confidence interval. As in Fig. 3, the left-most plot shows the overall result  
309 while the other three panels show results from individual major discipline. Reporting sex is  
310 associated with lower impact factors and the effect remains stable over time.

311  
312

313 **DISCUSSION**

314

315 Our results show that, over the last forty years, there has been a dramatic rise in sex-related  
316 reporting in clinical medicine and public health; yet, there has not been a concomitant rise of  
317 sex-related reporting in biomedical research, where only 31% of papers reported on sex in 2016.  
318 For clinical medicine and public health, percentages of sex-related reporting reached 67% and  
319 69% (respectively) in 2016. This confirms trends which have implied increasing rates of sex-  
320 related reporting (45); however, this is the first study to provide a proportion of the literature  
321 which is inclusive of all disciplines and specialities.

322

323 Our results demonstrated strong variation in sex-related reporting across disciplines. Some of  
324 these differences may seem intuitive: e.g., it is perhaps unsurprising that women are studied most  
325 often in gynecology. However, some of these imbalances can lead to grave consequences. Bias  
326 with regards to fertility studies has created a dangerous double standard in some clinical trials in  
327 which women must have contraceptive requirements, but men do not, even when paternal drug  
328 exposure may lead to fetal harm (19). Sex-related reporting is the first step towards improving  
329 ethical standards of research in regards to sex.

330

331 Area of research is only one factor that affects sex-related reporting in medical research. Papers  
332 with female first and last authors are more likely to report sex—especially female or both  
333 sexes—controlling for number of authors, representation of women in diseases, specialties,  
334 countries, continents, and publication years. These results complement recent results (18), which,  
335 based on the GenderMedDB (44), have shown that female first and last authors were more likely  
336 to report on sex. However, our results are based on a larger dataset— 3,394 vs 1.1 million papers  
337 reporting sex analyzed in the regressions—, with more controls and distinguishing between in  
338 the sex that is reported (female, male, or both). That is, while previous research has shown that  
339 female authors were more likely to report on sex, it did not demonstrate that women were also  
340 more likely to study females. This is a novel contribution of the present analysis.

341

342 Our analysis also provides evidence that research with sex-related reporting is more likely to  
343 appear in lower impact journals. Given their higher visibility, one might argue that high impact  
344 journals have a particular responsibility to enforce sex-related reporting when warranted.  
345 Furthermore, our regional analyses demonstrated that North American had the poorest rates of  
346 sex-related reporting across regions. This finding suggests that North American institutions must  
347 be proactive in order to achieve higher proportions of sex-related reporting in medical research.  
348 Analysis of sex-related reporting—at the journals, institutional, or country level—would be  
349 facilitated by greater standardization reporting practices in bibliographic indexes, which would  
350 lead to increased transparency.

351

352 **Limitations**

353

354 The use of indicators to measure science comes with some inherent limitations. We use MeSH as  
355 indicators of sex-related reporting in research. Our validation suggests that this approach is  
356 relatively accurate at identifying sex reporting, but is inadequate to analyze sex analysis. Further  
357 developments are necessary to ensure that sex-related data are provided to publishers and  
358 indexers in a nuanced and valid way for future analyses.

359  
360 We used journal-level classifications to indicate disciplines and specialities, based on the  
361 National Science Foundation classification. While this is standard in bibliometric analyses, it has  
362 limitations in the identification of papers' specific topics as well as potential misclassification of  
363 multidisciplinary research. The bibliometric alternative is the construction of a paper-level  
364 classification, but this comes with strong limitations, such as the lack of meaningful analytic  
365 clusters and the instability of clusters for diachronic analyses (45). We account for this limitation  
366 by taking diseases into account in our model.

367  
368 There are limitations to the use of authors' names as an indicator of their gender. Compared to  
369 self-report data, gender disambiguation algorithms are limited in that they can only be applied to  
370 those who have a full first name (rather than initials) and have a name that can be classified in a  
371 gender-binary way. There is therefore a sizeable proportion of authors of papers analyzed for  
372 which we could not assign gender, and this proportion varies by country, with a higher share of  
373 unassigned names in Asian countries.

374  
375 In our regression models, we did not explicitly model the missingness of the gender variables,  
376 adopting the ignorability assumption, as in a similar previous study (18). If the missingness of  
377 gender variables is strongly affected by unobserved factors, it may have produced biases in our  
378 results. Also, like in the aforementioned study (18), our main models also ignore the papers that  
379 do not have the disease MeSH terms with associated average female first (last) author fraction,  
380 although we note that the models that include such papers and do not use *f\_mesh* produce  
381 qualitatively similar results. The impact factor models have similar limitations. The relationship  
382 between the prestige of a journal and coverage of certain diseases associated with sex-related  
383 reporting should be taken into account in interpretation.

## 384 385 **CONCLUSION**

386  
387 At the cellular level—especially in the case of *in vitro* research using transformed cell lines—  
388 many researchers are simply unaware of the sex of the cell line they are using, despite efforts to  
389 document these cell lines (20). Although the process of creating stable and immortalized stem  
390 lines does not presently allow for perfect equivalency (leading to comparison) of female and  
391 male cell lines at this time, sex identification is nonetheless an important first step (21). This  
392 work is still in its infancy, but a full catalog of sex of common cell lines could increase the  
393 accuracy and degree of reporting. Science policy—from institutional to federal levels—should  
394 insist upon sex-related reporting for these studies.

395  
396 It is laudable that the NIH has achieved parity in terms of inclusion of females in clinical and  
397 health-based studies (9). Parity at the aggregate level, however, may obscure some differences at  
398 the field level. For example, our data results that females are more often studied in virology and  
399 cancer while males are the focus in neurology and the study of addictive diseases; these  
400 disparities may cause distortions in what is known about each sex within these fields. Research  
401 that examines both sexes extends the generalizability of the research, reduces the risk of practical  
402 health-based interventions and applications, and enhances replicability. It is important that parity  
403 be demonstrated at lower levels of analyses to mitigate disparities, particularly in specialties with  
404 implications for both sexes.

405  
406 When working with animal models, many researchers have simply used males as a default  
407 model; and the current generation has simply followed tradition. Given the growing importance  
408 of animal welfare, Institutional Animal Care and Use Committees (IACUC) ensure validity of  
409 research while also promoting the “three Rs”: replacement (with non-animals, e.g., cells, tissue),  
410 refinement (reduction of pain, suffering and distress) and reduction (in the number of animals)  
411 (22). If sex inclusion is not properly justified from the onset in the research design, reduction of  
412 the sample may make the population base too small for extensive sex stratification. This  
413 reinforces the relationship between sex-related reporting and research design: sex inclusion is  
414 more feasible when planned at the onset during research design.

415  
416 Sex inclusion is also a matter of scientific integrity. For example, Responsible Conduct of  
417 Research (RCR) training, which is obligatory for all publicly funded researchers in the US,  
418 examines issues of gender discrimination respecting scientists, and the inclusion of females in  
419 research on human subjects (e.g., clinical trials) (23-24). However, sex inclusion and reporting  
420 can and should be discussed in many other areas of research integrity. For example, micro-ethics  
421 discussions—often called “good laboratory practice” —should enable sex identification in  
422 effective record keeping, transparent reporting, and any sharing of data or material (such as on  
423 Material Transfer Agreements). Sex identification becomes an identifying factor that augments  
424 reproducibility and replicability. Research that considers sex differences could ultimately reduce  
425 health inequities, making sex-related reporting an ethical obligation and social responsibility.

426  
427 Journals have taken initial steps to adopt guidelines for reporting on sex-related reporting and  
428 analysis (15). However, this is more the exception rather than the rule. In 2011, the Institute of  
429 Medicine hosted a workshop on sex-related reporting in research with various stakeholders  
430 including editors in biomedical research and medicine. Editors and stakeholders agreed that sex-  
431 related reporting is feasible and fairly simple; however, requiring comparison between sexes—  
432 such as sex stratification—seemed controversial; many thought that controlling how experiments  
433 were designed, planned and conducted should be enabled and enforced mainly by funding  
434 agencies (25).

435  
436 One may hypothesize that since women are not prevalent in leadership positions, their presence  
437 may also be limited as editors, making sex-related reporting systemically less important to lead  
438 editors. It may also be that female authors have a limited ability to direct research within a lab:  
439 women hold a minority of authorships across the sciences (26), account for only a third of first-  
440 authorships in high impact medical journals (27), and are more likely to be involved in  
441 experimentation than in research design (38). Gender is also a factor in grant receipt and amount  
442 of funding (28). Without women leading and designing research, there may be markedly fewer  
443 articles with sex-inclusion generally, and studies of women, specifically. This has potentially  
444 dramatic health consequences for the entire population.

445  
446 Medical education, healthcare procurement, and service provision are expected to be based on  
447 the use of the best available scientific evidence. Therefore, the intentional or unintentional  
448 inclusion of sex biases “upstream” in research can be particularly pernicious for the  
449 “downstream” policy-making and service provision and policy. Sex and gender must be taken  
450 into account throughout the lifecycle of research. Diversification in the scientific workforce and

451 in the research populations—from cell lines, to rodents, to humans—is essential to produce the  
452 most rigorous and effective medical research.

453

## 454 **DECLARATIONS**

### 455 **Contributions**

456 Conceived and designed the experiments: CRS, ES, YYA, VL; Performed the experiments: CRS,  
457 YYA, BM, VL; Analyzed the data: CRS, YYA, VL; Wrote the paper: CRS, YYA, ES, VL. All  
458 authors approved the final text.

### 459 **Acknowledgements**

460 This study was funded by the Canada Research Chairs program (grant #950-231768). Authors  
461 thank David Endicott at the IU Statistical Consulting Center for helpful suggestions for the  
462 statistical analysis.

## 463 **REFERENCES**

- 464 1. B.M. Deasy et al.; A role for cell sex in stem cell-mediated skeletal muscle regeneration:  
465 female cells have higher muscle regeneration efficiency. *J. Cell. Bio.* 177, 73-86 (2007).
- 466 2. W.D. Nelson et al.; Sex-dependent attenuation of plague growth after treatment with bone  
467 marrow mononuclear cells. *Circ. Res.* 101, 1319-1327. (2007).
- 468 3. L. Du et al.; Starving neurons show sex difference in autophagy. *J. Biol. Chem.* 284,  
469 2383-2396 (2009).
- 470 4. V. Regitz-Zagrosek, Sex and gender differences in health. *EMBO Rep.* 13, 596-603  
471 (2012).
- 472 5. P.A. McCombe et al., Sexual dimorphism in autoimmune disease. *Curr. Mol. Med.* 9,  
473 1058-1079 (2018).
- 474 6. F. Franconi et al.; Gender differences in drug responses. *Pharmacol. Res.* 55, 81-95  
475 (2007).
- 476 7. N. Jochmann et al.; Female-specific aspects in the pharmacotherapy of chronic  
477 cardiovascular diseases. *Eur. Heart. J.* 26, 1585-95 (2005).
- 478 8. US General Accounting Office, Drug safety: most drugs withdrawn in recent years had  
479 greater health risks for women. *Gov. Publ. Office.* (2001).
- 480 9. J.A. Clayton, F.S. Collins, Policy: NIH to balance sex in cell and animal studies. *Nature.*  
481 509, 282-283 (2014).
- 482 10. A.K. Beery, I. Zucker, Sex bias in neuroscience and biomedical research. *Neurosci.*  
483 *Biobehav. Res.* 35, 5650572 (2011).
- 484 11. K.E. Taylor et al.; Reporting of sex as a variable in cardiovascular studies using cultured  
485 cells. *Biol. Sex. Differ.* 2, 11 (2011).
- 486 12. D.Y. Yoon et al.; Sex bias exists in basic science and translational surgical research.  
487 *Surgery.* 156, 508-516 (2014).
- 488 13. S.L. Klein et al.; Opinion: sex inclusion in basic research drives discovery. *Proc. Natl.*  
489 *Acad. Sci. U.S.A.* 112, 5257-5258 (2015).
- 490 14. B.J. Predergast, K.C. Onishi, I. Zucker, Female mice liberated for inclusion in  
491 neuroscience and biomedical research. *Neurosci. Biobehav. Res.* 40, 1-5 (2014).

- 492 15. C. Wald, C. Wu, Of mice and women: the bias in animal models. *Science*. 327, 1571-  
493 1572 (2010).
- 494 16. R.C. Rabin, Labs are told to start including a neglected variable: females. *New York*  
495 *Times*. May 15 (2014).
- 496 17. R.E. Sorge et al.; Olfactory exposure to males, including men, causes stress and related  
497 analgesia in rodents. *Nat. Methods*. 11, 629-632.
- 498 18. M.W. Nielsen, J.P. Andersen, L. Schiebinger, J.W. Schneider, One and a half million  
499 medical papers reveal a link between author gender and attention to gender and sex  
500 analysis. *Nat. Hum. Behav.* 1, 791-796.
- 501 19. R.J. DeLap, J.L. Fourcroy, G.A. Fleming. Fetal Harm Due to Paternal Drug Exposure: A  
502 Potential Issue in Drug Development. *Drug Information Journal* 30 (2): 359–64. (1996).
- 503 20. K. Shah, C.E. McCormach, N.A. Bradbury, Do you know the sex of your cells? *Am. J.*  
504 *Physiol. Cell. Physiol.* 306, C3-C18 (2014).
- 505 21. S.A. Ritz, et al. First Steps for Integrating Sex and Gender Considerations into Basic  
506 Experimental Biomedical Research. *The FASEB Journal* 28 (1): 4–13 (2014).
- 507 22. M. H. Lloyd, B.W. Foden, S.E. Wolfensohn. Refinement: Promoting the Three Rs in  
508 Practice. *Laboratory Animals* 42 (3): 284–93 (2008).
- 509 23. A.E. Shamoo, D.B. Resnik. Responsible Conduct of Research. Second Edition. *Oxford*  
510 *University Press*. 68-81pp. 236-289pp. (2015)
- 511 24. D.B. Resnik, The Ethics of Research with Human Subjects: Protecting People,  
512 Advancing Science, Promoting Trust. *Springer Nature*. 215-234 pp. (2018).
- 513 25. Institute of Medicine. Sex-Specific Reporting of Scientific Research: A Workshop  
514 Summary. Washington, DC: The National Academies Press.  
515 <https://doi.org/10.17226/13307> (2012). Last accessed Feb. 25 (2018).
- 516 26. V. Lariviere, Y., Gingras, B., Cronin, C.R. Sugimoto, Bibliometrics: global gender  
517 disparities in science. *Nature*. 504, 211-213 (2013).
- 518 27. G. Filardo et al.; Trends and comparison of female first authorship in high impact  
519 medical journals: observational study (1994-2014). *BMJ*. 352, i847 (2016).
- 520 28. R. Jagsi et al.; Similarities and differences in the career trajectories of male and female  
521 career development award recipients. *Acad. Med.* 86, 1315-1421 (2011).
- 522 29. V. Lariviere et al.; Contributorship and division of labor in knowledge production. *Soc.*  
523 *Stud. Sci.* 46, 417-431 (2016).
- 524 30. L. Doyal. Sex, gender, and health: the need for a new approach. *BMJ*. 323, 1061-63  
525 (2001).
- 526 31. R. Connell. Gender, health and theory: Conceptualizing the issue, in local and world  
527 perspective. *Soc. Sci. & Med.* 74, 1675-83. (2012).
- 528 32. C.C. Whitacre, S.C. Reingold, P.A. O’Looney. A gender gap in autoimmunity. *Science*.  
529 282, 1277-78. (1999).
- 530 33. S. Oertelt-Prigoione. The influence of sex and gender on the immune response.  
531 *Autoimmunity Rev.* A479-85. (2012).
- 532 34. S. Seabold, J. Perktold, Statsmodels: econometrics and statistical modeling with python.  
533 *Proc. 9th Python in Sci. Conf.* (2010).
- 534 35. S. Does et al., Implications of research staff demographics for psychological science.  
535 *American Psychologist*, 73, 639-650. (2006).

- 536 36. D.L. Lorenzetti, Y. Lin. Locating sex-and gender-specific data in health promotion  
537 research: evaluating the sensitivity and precision of published filters. *JMLA*, 105, 216  
538 (2017).
- 539 37. C.J. Moerman, R. Deurenberg, J.A. Haafkens. Locating sex-specific evidence on clinical  
540 questions in MEDLINE: a search filter for use on OvidSP™. *BMC medical research*  
541 *methodology*, 9(1), 25. (2009).
- 542 38. B. Macaluso, V. Larivière, T.J. Sugimoto, C.R. Sugimoto. Is science built on the  
543 shoulders of women? A study of gender differences in contributorship. *Acad Med*, 91,  
544 1136-42. (2016).
- 545 39. J. Johnson, Z. Sharman, B. Vissandjee, D.E. Stewart. Does a change in health research  
546 funding policy related to the integration of sex and gender have an impact?. *PLoS One*,  
547 9(6), e99900. (2014)
- 548 40. P. Allotey, C. Allotey-Reidpath, D.D Reidpath. Gender bias in clinical case reports: A  
549 cross-sectional study of the “big five” medical journals. *PLOS ONE*. 12, e0177386.  
550 (2017).
- 551 41. V. Larivière. C.R. Sugimoto. The Journal Impact Factor: A brief history, critique, and  
552 discussion of adverse effects. *arXiv:1801.08992*. (2018).
- 553 42. L. Santamaría, H. Mihaljević. Comparison and benchmark of name-to-gender inference  
554 services. *PeerJ Computer Science* 4:e156 (2018).
- 555 43. F. Karimi et al. Inferring gender from names on the web: A comparative evaluation of  
556 gender detection methods. *WWW’16 Companion*, pp. 53-54 (2016).
- 557 44. S. Oertelt-Prigione, B.O. Gohlke, M. Dunkel, R. Preissner, V. Regitz-Zagrosek.  
558 GenderMedDB: an interactive database of sex and gender-specific medical literature.  
559 *Biology of sex differences*, 5, 7. (2014).
- 560 45. C. R. Sugimoto, V. Larivière. *Measuring Research : what everyone needs to know*.  
561 Oxford University Press (2018).
- 562 46. S. Oertelt-Prigione, R. Parol, S. Krohn, R. Preißner, V. Regitz-Zagrosek (2010). Analysis  
563 of sex and gender-specific research reveals a common increase in publications and  
564 marked differences between disciplines. *BMC medicine*, 8, 70.  
565

566 **APPENDIX**

567

568 **MeSH analysis**

569 To test our method for false negatives, we used a specialties-based stratified sampling from all  
 570 articles that did *not* include a sex-related MeSH (Table S1). From these, we randomly selected  
 571 three articles for each specialty. In total, 171 articles (from all 57 specialties) were coded. For  
 572 147 of these studies (86%), neither sex nor gender was mentioned anywhere in the article.  
 573 Within those that did not mention sex, 43% (n=73) were cases where sex-based analysis was not  
 574 particularly warranted: e.g., they were largely non-empirical (reviews, policy papers, opinions)  
 575 or focused on computational or mathematical models. The remaining were empirical pieces  
 576 involving humans, animals, and cell cultures, in which sex-based reporting and analysis would  
 577 be expected. Only 14 studied explicitly reported the sex of the study: eight were single-sex  
 578 studies and six provided distributions of the sex of the population. Of the studies which explicitly  
 579 or implicitly mention sex, only four studies provided a sex-based analysis. In two of these,  
 580 although sex was controlled for in the regression, there was no distribution listed by sex.  
 581 Therefore, in only 1% of studies (two of 171) was sex both reported and analyzed. This confirms  
 582 the association between the lack of a sex-based MeSH and the absence of sex-based reporting  
 583 and analysis in the study (Table S1).

584

585 Table S1. Manual validation of sex-based MeSH headings

586

	Sex inclusion warranted	Sex reporting	Sex analysis
Absence of sex-related MeSH (false negatives) (n=171)	56% (n=96)	8% (n=14)	1% (n=2)
Presence of sex-related MeSH (false positives) (n=171)	99% (n=169)	95% (n=164)	76% (n=130)

587

588 We analyzed for false positives in a similar manner. A sample of 171 papers from the 57  
 589 disciplines were retrieved, with papers having male only, female only, and both sex-related  
 590 MeSH in equal proportion. Only two articles did not warrant sex-related reporting: one was  
 591 providing a blueprint for a genomic platform and the other a technical report on a medical  
 592 device. Out of the 169 remaining papers, 164 reported the sex of the population studied; the 5  
 593 studies that did not report sex of were cell-based analyses (n=3), a case study that did not report  
 594 the sex of the individual analyzed (n=1), and one empirical study. All misclassifications were  
 595 single-sex studies; all the papers to which two sex-related MeSH were assigned contained  
 596 information on the sex of the sample. However, while all single sex studies *de facto* reported and  
 597 analyzed findings by sex, this was not true of those where both sexes were in MeSH. For  
 598 instance, 33 of the 57 papers (58%) that had both sex-related MeSH assigned reported the sex  
 599 distribution of their sample, but did not break down the results or outcomes by sex. The  
 600 remainder (24 papers, 41%) contained both the sex breakdown of the sample as well as results  
 601 analyzed for each sex. This suggests that MeSH headings are good indicators of sex-related  
 602 reporting, but not sex analysis.



603 **Regression analysis.** A flowchart of the process is presented in Figure S1. In the sex-related  
604 reporting models, our dependent variables (SR, SR\_M, SR\_F, and SR\_B) are binary variables  
605 that indicates the existence of sex-related reporting in the paper, determined by the sex-related  
606 MeSH terms. In the main Sex-related reporting (SR) Model, ‘Male’, ‘Female’, and ‘Both’ are all  
607 considered to be one (SR=1) and ‘None’ is considered to be zero (SR=0); in SR\_B model, only  
608 ‘Both’ is considered to be one (SR\_B=1); in SR\_F model, ‘Female’ and ‘Both’ are considered to  
609 be one (SR\_F=1); SR\_M is analogously defined. To capture the general participation of female  
610 authors in disease topics, we calculated the average female first (last) author fraction given a set  
611 of disease MeSH terms (‘C’ category) in a paper. We first calculated the fraction of the papers  
612 with female first (last) author given a disease MeSH term, then for each paper that contains  
613 disease MeSH terms, we averaged the mean value associated with each MeSH term. Because the  
614 two variables for the first and the last authors are strongly correlated and may cause  
615 multicollinearity problem, we take the average of the two values to obtain the final variable  
616 *f\_mesh*. As in a similar study (18), we dropped articles for which we could not calculate MeSH  
617 covariates. The remaining dataset contains  $N = 1,273,687$  data points. The main results do not  
618 qualitatively change when we include all 2,018,697 articles without using *f\_mesh*. Similarly,  
619 using the main country label extracted from each paper, we obtain a country female author  
620 covariate, *f\_country*, based on the female first and last author prevalence in each country. The  
621 condition index of the design matrix was calculated to estimate the strength of the  
622 multicollinearity. After merging the two first- and last-author covariates, the condition index was  
623 smaller than 30 (26.1).

624  
625 We fitted logistic regression models using the standard enter method, with binary variables for  
626 the authors’ gender combinations—which has been associated with sex-related reporting—as the  
627 primary independent variables. We use the following control variables: the number of authors  
628 ( $\log_2$ ), binary variables for 57 specialties, binary variables for six continents (based on the  
629 affiliation of the author), average female author fraction for each disease MeSH term (*f\_mesh*),  
630 average female author fraction for each country (*f\_country*), and year. The number of authors  
631 reflect the scale of the study, which is likely associated with sex-related reporting. The  
632 specialties, continents, *f\_mesh*, and *f\_country* are included to control for the association between  
633 author's gender, topics, diseases, and sex-related reporting. To capture the effect of the authors’  
634 gender, we created four categorical dummy variables (‘MM’, ‘MF’, ‘FM’, ‘FF’) and used ‘MM’  
635 as the reference. The specialties and continents are similarly prepared. North America was used  
636 as the reference. When all specialties are considered, the reference variable is “Addictive  
637 Diseases”; For disciplinary models, the reference variables are “Anatomy & Morphology”,  
638 “Addictive Diseases” “Geriatrics & Gerontology” for Biomedical Research, Clinical Medicine,  
639 and Public Health respectively. The number of authors exhibits a heavy-tailed distribution that  
640 spans multiple orders of magnitude. Therefore, we used a logarithmic transformation with base 2  
641 instead of using the raw values. The regression tables are shown in Tables S3-S6 (SR models for  
642 the whole dataset and for three major disciplines) and SI Tables S7-S9 (SR\_M, SR\_F, SR\_B  
643 models for the whole dataset).

644 In the impact factor model, we used OLS (Ordinary Least Square) linear regression model with  
645 the following independent variables: binary variables for sex-related reporting (Male, Female,  
646 Both), and control variables: binary variables for the gender of the first and the last authors, the  
647 number of authors ( $\log_2$ ), binary variables for the specialties and continents. In order to examine  
648 temporal trends, we extracted papers published in 2008, 2010, 2012, 2014, and 2016, and fitted

649 four models for each year set. As in the case of the previous model, we fitted an aggregated  
 650 model that includes all major disciplines as well as separate discipline-based models (see Tables  
 651 S10-S13). Multicollinearity is tested using the variance inflation factor (VIF) and none of our  
 652 independent variables exhibits high (>5.0) VIF. All models were estimated using Python  
 653 statsmodels package (34) and source code is available on GitHub (<https://github.com/TBD>).

654  
 655 **Additional tables and figures**

656  
 657 **Table S2.** Sex-related MeSH, number of papers retrieved, and percentage of all papers retrieved,  
 658 1980-2016

Female-related MeSH	Male-related MeSH
Battered Women	Circumcision, Male
Circumcision, Female	Contraceptive Agents, Male
Condoms, Female	Contraceptive Devices, Male
Contraceptive Agents, Female	Fertility Agents, Male
Contraceptive Devices, Female	Genital Diseases, Male
Dentists, Women	Genital Neoplasms, Male
Female	Genitalia, Male
Female Athlete Triad Syndrome	Homosexuality, Male
Female Urogenital Diseases	Infertility, Male
Fertility Agents, Female	Male
Genital Diseases, Female	Men's Health
Genital Neoplasms, Female	Nurses, Male
Genitalia, Female	Tuberculosis, Male Genital
Homosexuality, Female	Urologic Surgical Procedures, Male
Infertility, Female	
Physicians, Women	
Pregnant Women	
Tuberculosis, Female Genital	
Women	
Women, Working	
Women's Health	
Women's Health Services	
Women's Rights	

660

661

662 **Table S3.** Coefficients and odds ratios of the variables from the logistic regression model for the  
 663 sex-related reporting on the aggregated dataset with all disciplines and specialties

Logit Regression Results

Dep. Variable:	SR	Log-Likelihood:	6.9E+05
Model:	Logit	LL-Null:	7.8E+05
Method:	MLE	LLR p-value:	0.0E+00
No. Observations:	1273687		
Df Residuals:	1273618		
Df Model:	68		
Pseudo R-sq.	0.1167		

	coef	std err	z	P> z	[0.025	0.975]	Odds Ratio	[0.025	0.975]
Intercept	5.2113	1.714	3.04	0.002	1.852	8.571	183.33	6.37	5276.4
Male-Female	0.0597	0.007	9.119	0	0.047	0.072	1.06	1.05	1.07
Female-Male	0.063	0.005	12.215	0	0.053	0.073	1.07	1.05	1.08
Female-Female	0.2313	0.007	34.423	0	0.218	0.245	1.26	1.24	1.28
CONTINENT[T.Africa]	1.0697	0.028	37.756	0	1.014	1.125	2.91	2.76	3.08
CONTINENT[T.Asia]	0.4913	0.007	73.224	0	0.478	0.504	1.63	1.61	1.66
CONTINENT[T.Europe]	0.1412	0.005	27.639	0	0.131	0.151	1.15	1.14	1.16
CONTINENT[T.Oceania]	0.042	0.013	3.346	0.001	0.017	0.067	1.04	1.02	1.07
CONTINENT[T.South America]	0.317	0.013	24.151	0	0.291	0.343	1.37	1.34	1.41
SUBDISCIPLINE[T.Allergy]	-1.3039	0.052	-25.246	0	-1.405	-1.203	0.27	0.25	0.3
SUBDISCIPLINE[T.Anatomy & Morphology]	-1.1473	0.07	-16.371	0	-1.285	-1.01	0.32	0.28	0.36
SUBDISCIPLINE[T.Anesthesiology]	-0.7386	0.043	-17.184	0	-0.823	-0.654	0.48	0.44	0.52
SUBDISCIPLINE[T.Arthritis & Rheumatology]	-0.884	0.037	-23.937	0	-0.956	-0.812	0.41	0.38	0.44
SUBDISCIPLINE[T.Biochemistry & Molecular Biology]	-2.6924	0.033	-82.787	0	-2.756	-2.629	0.07	0.06	0.07
SUBDISCIPLINE[T.Biomedical Engineering]	-2.5512	0.042	-60.833	0	-2.633	-2.469	0.08	0.07	0.08
SUBDISCIPLINE[T.Biomedical Social Sciences]	-0.6888	0.077	-9.003	0	-0.839	-0.539	0.5	0.43	0.58
SUBDISCIPLINE[T.Biophysics]	-2.37	0.059	-40.492	0	-2.485	-2.255	0.09	0.08	0.1
SUBDISCIPLINE[T.Cancer]	-1.3226	0.032	-41.57	0	-1.385	-1.26	0.27	0.25	0.28
SUBDISCIPLINE[T.Cardiovascular System]	-0.7492	0.033	-22.966	0	-0.813	-0.685	0.47	0.44	0.5

SUBDISCIPLINE[T.Cellular Biology Cytology & Histology]	-2.5332	0.035	-71.793	0	-2.602	-2.464	0.08	0.07	0.09
SUBDISCIPLINE[T.Dentistry]	-0.957	0.036	-26.252	0	-1.028	-0.886	0.38	0.36	0.41
SUBDISCIPLINE[T.Dermatology & Venereal Disease]	-0.7223	0.035	-20.4	0	-0.792	-0.653	0.49	0.45	0.52
SUBDISCIPLINE[T.Embryology]	-1.3131	0.063	-20.887	0	-1.436	-1.19	0.27	0.24	0.3
SUBDISCIPLINE[T.Endocrinology]	-0.784	0.034	-22.976	0	-0.851	-0.717	0.46	0.43	0.49
SUBDISCIPLINE[T.Environmental & Occupational Health]	-0.6268	0.039	-16.221	0	-0.703	-0.551	0.53	0.5	0.58
SUBDISCIPLINE[T.Fertility]	2.2613	0.124	18.164	0	2.017	2.505	9.6	7.52	12.24
SUBDISCIPLINE[T.Gastroenterology]	-1.2015	0.033	-35.915	0	-1.267	-1.136	0.3	0.28	0.32
SUBDISCIPLINE[T.General & Internal Medicine]	-1.0613	0.032	-33.184	0	-1.124	-0.999	0.35	0.32	0.37
SUBDISCIPLINE[T.General Biomedical Research]	-2.0325	0.032	-63.209	0	-2.096	-1.97	0.13	0.12	0.14
SUBDISCIPLINE[T.Genetics & Heredity]	-2.0637	0.034	-60.237	0	-2.131	-1.997	0.13	0.12	0.14
SUBDISCIPLINE[T.Geriatrics]	-0.3942	0.051	-7.739	0	-0.494	-0.294	0.67	0.61	0.75
SUBDISCIPLINE[T.Geriatrics & Gerontology]	-0.8672	0.048	-17.905	0	-0.962	-0.772	0.42	0.38	0.46
SUBDISCIPLINE[T.Health Policy & Services]	-0.9314	0.038	-24.609	0	-1.006	-0.857	0.39	0.37	0.42
SUBDISCIPLINE[T.Hematology]	-1.6181	0.035	-46.584	0	-1.686	-1.55	0.2	0.19	0.21
SUBDISCIPLINE[T.Immunology]	-1.9593	0.032	-60.81	0	-2.022	-1.896	0.14	0.13	0.15
SUBDISCIPLINE[T.Microbiology]	-2.7061	0.036	-76.111	0	-2.776	-2.636	0.07	0.06	0.07
SUBDISCIPLINE[T.Microscopy]	-2.9934	0.165	-18.109	0	-3.317	-2.669	0.05	0.04	0.07
SUBDISCIPLINE[T.Miscellaneous Biomedical Research]	-1.6276	0.04	-40.223	0	-1.707	-1.548	0.2	0.18	0.21
SUBDISCIPLINE[T.Miscellaneous Clinical Medicine]	-0.8886	0.037	-24.09	0	-0.961	-0.816	0.41	0.38	0.44
SUBDISCIPLINE[T.Nephrology]	-1.0839	0.039	-27.559	0	-1.161	-1.007	0.34	0.31	0.37
SUBDISCIPLINE[T.Neurology & Neurosurgery]	-0.8158	0.032	-25.557	0	-0.878	-0.753	0.44	0.42	0.47

SUBDISCIPLINE[T.Nursing]	-0.802	0.037	-21.424	0	-0.875	-0.729	0.45	0.42	0.48
SUBDISCIPLINE[T.Nutrition & Dietetic]	-0.9701	0.037	-26.093	0	-1.043	-0.897	0.38	0.35	0.41
SUBDISCIPLINE[T.Obstetrics & Gynecology]	1.6887	0.055	30.768	0	1.581	1.796	5.41	4.86	6.03
SUBDISCIPLINE[T.Ophthalmology]	-0.615	0.035	-17.645	0	-0.683	-0.547	0.54	0.51	0.58
SUBDISCIPLINE[T.Orthopedics]	-0.4655	0.035	-13.427	0	-0.533	-0.398	0.63	0.59	0.67
SUBDISCIPLINE[T.Otorhinolaryngology]	-0.4285	0.036	-11.793	0	-0.5	-0.357	0.65	0.61	0.7
SUBDISCIPLINE[T.Parasitology]	-2.5952	0.041	-63.308	0	-2.676	-2.515	0.07	0.07	0.08
SUBDISCIPLINE[T.Pathology]	-1.1201	0.034	-32.53	0	-1.188	-1.053	0.33	0.3	0.35
SUBDISCIPLINE[T.Pediatrics]	-0.5595	0.035	-16.022	0	-0.628	-0.491	0.57	0.53	0.61
SUBDISCIPLINE[T.Pharmacology]	-1.9038	0.032	-59.283	0	-1.967	-1.841	0.15	0.14	0.16
SUBDISCIPLINE[T.Pharmacy]	-2.2358	0.048	-46.92	0	-2.329	-2.142	0.11	0.1	0.12
SUBDISCIPLINE[T.Physiology]	-1.1413	0.037	-30.479	0	-1.215	-1.068	0.32	0.3	0.34
SUBDISCIPLINE[T.Psychiatry]	-0.0969	0.038	-2.553	0.011	-0.171	-0.023	0.91	0.84	0.98
SUBDISCIPLINE[T.Public Health]	-0.587	0.035	-16.571	0	-0.656	-0.518	0.56	0.52	0.6
SUBDISCIPLINE[T.Radiology & Nuclear Medicine]	-0.8194	0.034	-24.41	0	-0.885	-0.754	0.44	0.41	0.47
SUBDISCIPLINE[T.Rehabilitation]	-0.2948	0.041	-7.131	0	-0.376	-0.214	0.74	0.69	0.81
SUBDISCIPLINE[T.Respiratory System]	-1.0012	0.036	-27.573	0	-1.072	-0.93	0.37	0.34	0.39
SUBDISCIPLINE[T.Social Studies of Medicine]	-1.5732	0.225	-6.983	0	-2.015	-1.132	0.21	0.13	0.32
SUBDISCIPLINE[T.Speech-Language Pathology and Audiology]	-0.3589	0.062	-5.751	0	-0.481	-0.237	0.7	0.62	0.79
SUBDISCIPLINE[T.Surgery]	-0.3288	0.033	-10.036	0	-0.393	-0.265	0.72	0.68	0.77
SUBDISCIPLINE[T.Tropical Medicine]	-1.691	0.043	-39.603	0	-1.775	-1.607	0.18	0.17	0.2
SUBDISCIPLINE[T.Urology]	-0.0302	0.037	-0.828	0.408	-0.102	0.041	0.97	0.9	1.04

SUBDISCIPLINE[T.Veterinary Medicine]	-1.7557	0.035	-50.708	0	-1.824	-1.688	0.17	0.16	0.18
SUBDISCIPLINE[T.Virology]	-2.8151	0.036	-77.147	0	-2.887	-2.744	0.06	0.06	0.06
YEAR	-0.0028	0.001	-3.235	0.001	-0.004	-0.001	1	1	1
np.log2(N_AUTHORS)	0.6706	0.003	233.879	0	0.665	0.676	1.96	1.94	1.97
F_MESH	0.8641	0.034	25.329	0	0.797	0.931	2.37	2.22	2.54
F_COUNTRY	0.7529	0.039	19.185	0	0.676	0.83	2.12	1.97	2.29

664

665 **Table S4.** Coefficients and odds ratios of the variables from the logistic regression on  
666 Biomedical Research.

Logit Regression Results

Dep. Variable:	SR	Log-Likelihood:	1.4E+05
Model:	Logit	LL-Null:	1.5E+05
Method:	MLE	LLR p-value:	0.0E+00
No. Observations:	219215		
Df Residuals:	219188		
Df Model:	26		
	0.0847		
Pseudo R-sq.	1		

	coef	std err	z	P> z	[0.025	0.975]	Odds Ratio	[0.025	0.975]
Intercept	-50.969	3.931	-12.967	0	-58.673	-43.265	0	0	0
Male-Female	0.0706	0.014	4.963	0	0.043	0.098	1.07	1.04	1.1
Female-Male	0.0702	0.011	6.339	0	0.049	0.092	1.07	1.05	1.1
Female-Female	0.2369	0.014	16.799	0	0.209	0.264	1.27	1.23	1.3
CONTINENT[T.Africa]	1.1695	0.058	20.068	0	1.055	1.284	3.22	2.87	3.61
CONTINENT[T.Asia]	0.5853	0.014	42.402	0	0.558	0.612	1.8	1.75	1.84
CONTINENT[T.Europe]	0.2308	0.012	19.77	0	0.208	0.254	1.26	1.23	1.29
CONTINENT[T.Oceania]	0.1535	0.03	5.162	0	0.095	0.212	1.17	1.1	1.24
CONTINENT[T.South America]	0.475	0.028	16.805	0	0.42	0.53	1.61	1.52	1.7
SUBDISCIPLINE[T.Biochemistry & Molecular Biology]	-1.5374	0.064	-24.177	0	-1.662	-1.413	0.21	0.19	0.24
SUBDISCIPLINE[T.Biomedical Engineering]	-1.3926	0.069	-20.322	0	-1.527	-1.258	0.25	0.22	0.28
SUBDISCIPLINE[T.Biophysics]	-1.2211	0.08	-15.332	0	-1.377	-1.065	0.29	0.25	0.34
SUBDISCIPLINE[T.Cellular Biology Cytology & Histology]	-1.3836	0.065	-21.306	0	-1.511	-1.256	0.25	0.22	0.28
SUBDISCIPLINE[T.Embryology]	-0.2414	0.083	-2.893	0.004	-0.405	-0.078	0.79	0.67	0.92

SUBDISCIPLINE[T.General Biomedical Research]	-0.8852	0.063	-13.954	0	-1.01	-0.761	0.41	0.36	0.47
SUBDISCIPLINE[T.Genetics & Heredity]	-0.9187	0.065	-14.193	0	-1.046	-0.792	0.4	0.35	0.45
SUBDISCIPLINE[T.Microbiology]	-1.5937	0.065	-24.366	0	-1.722	-1.466	0.2	0.18	0.23
SUBDISCIPLINE[T.Microscopy]	-1.9419	0.174	-11.133	0	-2.284	-1.6	0.14	0.1	0.2
SUBDISCIPLINE[T.Miscellaneous Biomedical Research]	-0.5008	0.068	-7.394	0	-0.634	-0.368	0.61	0.53	0.69
SUBDISCIPLINE[T.Nutrition & Dietetic]	0.0956	0.066	1.44	0.15	-0.035	0.226	1.1	0.97	1.25
SUBDISCIPLINE[T.Parasitology]	-1.5184	0.069	-22.145	0	-1.653	-1.384	0.22	0.19	0.25
SUBDISCIPLINE[T.Physiology]	0.0145	0.066	0.22	0.826	-0.115	0.144	1.01	0.89	1.15
SUBDISCIPLINE[T.Virology]	-1.6637	0.066	-25.215	0	-1.793	-1.534	0.19	0.17	0.22
YEAR	0.0245	0.002	12.529	0	0.021	0.028	1.02	1.02	1.03
np.log2(N_AUTHORS)	0.537	0.006	87.617	0	0.525	0.549	1.71	1.69	1.73
F_MESH	1.5799	0.078	20.21	0	1.427	1.733	4.85	4.17	5.66
F_COUNTRY	1.4196	0.085	16.663	0	1.253	1.587	4.14	3.5	4.89

667

668 **Table S5.** Coefficients and odds ratios of the variables from the logistic regression on the  
669 **Clinical Medicine.**

Logit Regression Results

Dep. Variable:	SR	Log-Likelihood:	5.2E+05
Model:	Logit	LL-Null:	5.8E+05
Method:	MLE	LLR p-value:	0.0E+00
No. Observations:	995511		
Df Residuals:	995465		
Df Model:	45		
Pseudo R-sq.	0.09546		

	coef	std err	z	P> z	[0.025	0.975]	Odds Ratio	[0.025	0.975]
Intercept	21.6663	1.963	11.036	0	17.818	25.514	2.6E+09	5.5E+07	1.2E+11
Male-Female	0.0595	0.008	7.853	0	0.045	0.074	1.06	1.05	1.08
Female-Male	0.058	0.006	9.722	0	0.046	0.07	1.06	1.05	1.07
Female-Female	0.2318	0.008	28.882	0	0.216	0.247	1.26	1.24	1.28
CONTINENT[T.Africa]	1.0635	0.033	31.954	0	0.998	1.129	2.9	2.71	3.09
CONTINENT[T.Asia]	0.4521	0.008	57.646	0	0.437	0.467	1.57	1.55	1.6
CONTINENT[T.Europe]	0.1117	0.006	19.089	0	0.1	0.123	1.12	1.11	1.13
CONTINENT[T.Oceania]	0.036	0.015	2.453	0.014	0.007	0.065	1.04	1.01	1.07

CONTINENT[T.South America]	0.2951	0.015	19.244	0	0.265	0.325	1.34	1.3	1.38
SUBDISCIPLINE[T.Allergy]	-1.3333	0.052	-25.717	0	-1.435	-1.232	0.26	0.24	0.29
SUBDISCIPLINE[T.Anesthesiology]	-0.7808	0.043	-18.093	0	-0.865	-0.696	0.46	0.42	0.5
SUBDISCIPLINE[T.Arthritis & Rheumatology]	-0.9081	0.037	-24.501	0	-0.981	-0.835	0.4	0.37	0.43
SUBDISCIPLINE[T.Cancer]	-1.372	0.032	-42.946	0	-1.435	-1.309	0.25	0.24	0.27
SUBDISCIPLINE[T.Cardiovascular System]	-0.808	0.033	-24.616	0	-0.872	-0.744	0.45	0.42	0.48
SUBDISCIPLINE[T.Dentistry]	-0.9756	0.037	-26.659	0	-1.047	-0.904	0.38	0.35	0.4
SUBDISCIPLINE[T.Dermatology & Venereal Disease]	-0.723	0.035	-20.371	0	-0.793	-0.653	0.49	0.45	0.52
SUBDISCIPLINE[T.Endocrinology]	-0.8021	0.034	-23.442	0	-0.869	-0.735	0.45	0.42	0.48
SUBDISCIPLINE[T.Environmental & Occupational Health]	-0.6334	0.039	-16.352	0	-0.709	-0.557	0.53	0.49	0.57
SUBDISCIPLINE[T.Fertility]	2.2553	0.125	18.108	0	2.011	2.499	9.54	7.47	12.17
SUBDISCIPLINE[T.Gastroenterology]	-1.2516	0.034	-37.21	0	-1.318	-1.186	0.29	0.27	0.31
SUBDISCIPLINE[T.General & Internal Medicine]	-1.0823	0.032	-33.71	0	-1.145	-1.019	0.34	0.32	0.36
SUBDISCIPLINE[T.Geriatrics]	-0.4096	0.051	-8.021	0	-0.51	-0.31	0.66	0.6	0.73
SUBDISCIPLINE[T.Hematology]	-1.6656	0.035	-47.743	0	-1.734	-1.597	0.19	0.18	0.2
SUBDISCIPLINE[T.Immunology]	-1.9904	0.032	-61.581	0	-2.054	-1.927	0.14	0.13	0.15
SUBDISCIPLINE[T.Miscellaneous Clinical Medicine]	-0.9122	0.037	-24.651	0	-0.985	-0.84	0.4	0.37	0.43
SUBDISCIPLINE[T.Nephrology]	-1.1254	0.04	-28.489	0	-1.203	-1.048	0.32	0.3	0.35
SUBDISCIPLINE[T.Neurology & Neurosurgery]	-0.8478	0.032	-26.465	0	-0.911	-0.785	0.43	0.4	0.46
SUBDISCIPLINE[T.Obstetrics & Gynecology]	1.6966	0.055	30.881	0	1.589	1.804	5.46	4.9	6.07
SUBDISCIPLINE[T.Ophthalmology]	-0.6444	0.035	-18.42	0	-0.713	-0.576	0.52	0.49	0.56
SUBDISCIPLINE[T.Orthopedics]	-0.5266	0.035	-15.078	0	-0.595	-0.458	0.59	0.55	0.63
SUBDISCIPLINE[T.Otorhinolaryngology]	-0.4592	0.036	-12.591	0	-0.531	-0.388	0.63	0.59	0.68



SUBDISCIPLINE[T.Pathology]	-1.1565	0.035	-33.448	0	-1.224	-1.089	0.31	0.29	0.34
SUBDISCIPLINE[T.Pediatrics]	-0.5701	0.035	-16.287	0	-0.639	-0.502	0.57	0.53	0.61
SUBDISCIPLINE[T.Pharmacology]	-1.9301	0.032	-59.885	0	-1.993	-1.867	0.15	0.14	0.15
SUBDISCIPLINE[T.Pharmacy]	-2.2689	0.048	-47.443	0	-2.363	-2.175	0.1	0.09	0.11
SUBDISCIPLINE[T.Psychiatry]	-0.1007	0.038	-2.646	0.008	-0.175	-0.026	0.9	0.84	0.97
SUBDISCIPLINE[T.Radiology & Nuclear Medicine]	-0.8725	0.034	-25.869	0	-0.939	-0.806	0.42	0.39	0.45
SUBDISCIPLINE[T.Respiratory System]	-1.0311	0.036	-28.295	0	-1.102	-0.96	0.36	0.33	0.38
SUBDISCIPLINE[T.Surgery]	-0.3806	0.033	-11.554	0	-0.445	-0.316	0.68	0.64	0.73
SUBDISCIPLINE[T.Tropical Medicine]	-1.6946	0.043	-39.484	0	-1.779	-1.611	0.18	0.17	0.2
SUBDISCIPLINE[T.Urology]	-0.0755	0.037	-2.059	0.039	-0.147	-0.004	0.93	0.86	1
SUBDISCIPLINE[T.Veterinary Medicine]	-1.776	0.035	-51.152	0	-1.844	-1.708	0.17	0.16	0.18
YEAR	-0.0109	0.001	-11.142	0	-0.013	-0.009	0.99	0.99	0.99
np.log2(N_AUTHORS)	0.724	0.003	217.387	0	0.718	0.731	2.06	2.05	2.08
F_MESH	0.5421	0.04	13.663	0	0.464	0.62	1.72	1.59	1.86
F_COUNTRY	0.5278	0.045	11.726	0	0.44	0.616	1.7	1.55	1.85

670

671 **Table S6.** Coefficients and odds ratios of the variables from the logistic regression on the Public  
672 Health.

Logit Regression Results

Dep. Variable:	SR	Log-Likelihood:	3.2E+04
Model:	Logit	LL-Null:	3.3E+04
Method:	MLE	LLR p-value:	0.0E+00
No. Observations:	58961		
Df Residuals:	58941		
Df Model:	19		
Pseudo R-sq.	0.0317		9

	coef	std err	z	P> z	[0.025	0.975]	Odds Ratio	[0.025	0.975]
Intercept	-38.214	8.144	-4.692	0	-54.177	-22.252	0	0	0
Male-Female	0.0931	0.033	2.793	0.005	0.028	0.159	1.1	1.03	1.17
Female-Male	0.1956	0.028	6.948	0	0.14	0.251	1.22	1.15	1.29
Female-Female	0.2407	0.028	8.598	0	0.186	0.296	1.27	1.2	1.34

CONTINENT[T.Africa]	0.4562	0.125	3.645	0	0.211	0.701	1.58	1.23	2.02
CONTINENT[T.Asia]	0.8725	0.048	18.223	0	0.779	0.966	2.39	2.18	2.63
CONTINENT[T.Europe]	0.1096	0.025	4.331	0	0.06	0.159	1.12	1.06	1.17
CONTINENT[T.Oceania]	-0.1718	0.043	-3.968	0	-0.257	-0.087	0.84	0.77	0.92
CONTINENT[T.South America]	-0.1904	0.057	-3.341	0.001	-0.302	-0.079	0.83	0.74	0.92
SUBDISCIPLINE[T.Geriatrics & Gerontology]	-0.08	0.079	-1.011	0.312	-0.235	0.075	0.92	0.79	1.08
SUBDISCIPLINE[T.Health Policy & Services]	-0.0989	0.073	-1.35	0.177	-0.242	0.045	0.91	0.79	1.05
SUBDISCIPLINE[T.Nursing]	-0.0608	0.073	-0.832	0.406	-0.204	0.082	0.94	0.82	1.09
SUBDISCIPLINE[T.Public Health]	0.2299	0.072	3.188	0.001	0.089	0.371	1.26	1.09	1.45
SUBDISCIPLINE[T.Rehabilitation ]	0.4814	0.075	6.426	0	0.335	0.628	1.62	1.4	1.87
SUBDISCIPLINE[T.Social Studies of Medicine]	-0.8241	0.235	-3.512	0	-1.284	-0.364	0.44	0.28	0.69
SUBDISCIPLINE[T.Speech-Language Pathology and Audiology]	0.2632	0.089	2.971	0.003	0.09	0.437	1.3	1.09	1.55
YEAR	0.0182	0.004	4.504	0	0.01	0.026	1.02	1.01	1.03
np.log2(N_AUTHORS)	0.3914	0.015	26.813	0	0.363	0.42	1.48	1.44	1.52
F_MESH	2.1555	0.131	16.427	0	1.898	2.413	8.63	6.67	11.17
F_COUNTRY	1.8367	0.248	7.407	0	1.351	2.323	6.28	3.86	10.21

673

674 **Table S7.** Coefficients and odds ratios of the variables from the logistic regression, sex-related  
675 reporting = male (SR\_M).

Logit Regression Results

Dep. Variable:	SR_M	Log-Likelihood:	7.8E+05
Model:	Logit	LL-Null:	8.8E+05
Method:	MLE	LLR p-value:	0.0E+00
No. Observations:	1273687		
Df Residuals:	1273618		
Df Model:	68		
Pseudo R-sq.	0.1151		

	coef	std err	z	P> z	[0.025	0.975]	Odds Ratio	[0.025	0.975]
Intercept	-17.129	1.595	-10.737	0	-20.256	-14.002	0	0	0
Male-Female	0.0271	0.006	4.416	0	0.015	0.039	1.03	1.02	1.04
Female-Male	0.0114	0.005	2.36	0.018	0.002	0.021	1.01	1	1.02
Female-Female	0.0468	0.006	7.586	0	0.035	0.059	1.05	1.04	1.06
CONTINENT[T.Africa]	0.9095	0.024	37.802	0	0.862	0.957	2.48	2.37	2.6
CONTINENT[T.Asia]	0.3881	0.006	62.959	0	0.376	0.4	1.47	1.46	1.49

CONTINENT[T.Europe]	0.1589	0.005	33.075	0	0.15	0.168	1.17	1.16	1.18
CONTINENT[T.Oceania]	0.1179	0.012	9.972	0	0.095	0.141	1.13	1.1	1.15
CONTINENT[T.South America]	0.2494	0.012	21.133	0	0.226	0.272	1.28	1.25	1.31
SUBDISCIPLINE[T.Allergy]	-1.3335	0.048	-27.736	0	-1.428	-1.239	0.26	0.24	0.29
SUBDISCIPLINE[T.Anatomy & Morphology]	-1.9144	0.066	-29.062	0	-2.044	-1.785	0.15	0.13	0.17
SUBDISCIPLINE[T.Anesthesiology]	-1.5873	0.038	-41.674	0	-1.662	-1.513	0.2	0.19	0.22
SUBDISCIPLINE[T.Arthritis & Rheumatology]	-1.4279	0.032	-43.978	0	-1.491	-1.364	0.24	0.23	0.26
SUBDISCIPLINE[T.Biochemistry & Molecular Biology]	-3.0532	0.029	-104.68	0	-3.11	-2.996	0.05	0.04	0.05
SUBDISCIPLINE[T.Biomedical Engineering]	-3.1061	0.041	-76.173	0	-3.186	-3.026	0.04	0.04	0.05
SUBDISCIPLINE[T.Biomedical Social Sciences]	-0.9482	0.07	-13.527	0	-1.086	-0.811	0.39	0.34	0.44
SUBDISCIPLINE[T.Biophysics]	-2.9111	0.06	-48.841	0	-3.028	-2.794	0.05	0.05	0.06
SUBDISCIPLINE[T.Cancer]	-2.4319	0.028	-86.993	0	-2.487	-2.377	0.09	0.08	0.09
SUBDISCIPLINE[T.Cardiovascular System]	-1.3896	0.029	-48.477	0	-1.446	-1.333	0.25	0.24	0.26
SUBDISCIPLINE[T.Cellular Biology Cytology & Histology]	-3.0652	0.033	-93.343	0	-3.13	-3.001	0.05	0.04	0.05
SUBDISCIPLINE[T.Dentistry]	-1.4642	0.032	-45.224	0	-1.528	-1.401	0.23	0.22	0.25
SUBDISCIPLINE[T.Dermatology & Venereal Disease]	-1.3305	0.031	-43.054	0	-1.391	-1.27	0.26	0.25	0.28
SUBDISCIPLINE[T.Embryology]	-2.3549	0.059	-39.914	0	-2.471	-2.239	0.09	0.08	0.11
SUBDISCIPLINE[T.Endocrinology]	-1.3604	0.03	-45.782	0	-1.419	-1.302	0.26	0.24	0.27
SUBDISCIPLINE[T.Environmental & Occupational Health]	-0.9526	0.034	-28.272	0	-1.019	-0.887	0.39	0.36	0.41
SUBDISCIPLINE[T.Fertility]	-2.4977	0.041	-61.188	0	-2.578	-2.418	0.08	0.08	0.09
SUBDISCIPLINE[T.Gastroenterology]	-1.5788	0.03	-53.256	0	-1.637	-1.521	0.21	0.19	0.22
SUBDISCIPLINE[T.General & Internal Medicine]	-1.6525	0.028	-58.87	0	-1.708	-1.598	0.19	0.18	0.2
SUBDISCIPLINE[T.General Biomedical Research]	-2.415	0.028	-84.876	0	-2.471	-2.359	0.09	0.08	0.09

SUBDISCIPLINE[T.Genetics & Heredity]	-2.336	0.031	-75.676	0	-2.397	-2.276	0.1	0.09	0.1
SUBDISCIPLINE[T.Geriatrics]	-1.0436	0.043	-24.507	0	-1.127	-0.96	0.35	0.32	0.38
SUBDISCIPLINE[T.Geriatrics & Gerontology]	-0.9341	0.044	-21.029	0	-1.021	-0.847	0.39	0.36	0.43
SUBDISCIPLINE[T.Health Policy & Services]	-1.1178	0.034	-32.944	0	-1.184	-1.051	0.33	0.31	0.35
SUBDISCIPLINE[T.Hematology]	-1.8639	0.031	-60.049	0	-1.925	-1.803	0.16	0.15	0.16
SUBDISCIPLINE[T.Immunology]	-2.2844	0.029	-80.034	0	-2.34	-2.228	0.1	0.1	0.11
SUBDISCIPLINE[T.Microbiology]	-2.9776	0.033	-89.966	0	-3.043	-2.913	0.05	0.05	0.05
SUBDISCIPLINE[T.Microscopy]	-3.3816	0.187	-18.132	0	-3.747	-3.016	0.03	0.02	0.05
SUBDISCIPLINE[T.Miscellaneous Biomedical Research]	-2.2815	0.037	-61.008	0	-2.355	-2.208	0.1	0.09	0.11
SUBDISCIPLINE[T.Miscellaneous Clinical Medicine]	-1.3222	0.033	-40.173	0	-1.387	-1.258	0.27	0.25	0.28
SUBDISCIPLINE[T.Nephrology]	-1.4221	0.036	-40.026	0	-1.492	-1.352	0.24	0.22	0.26
SUBDISCIPLINE[T.Neurology & Neurosurgery]	-1.2039	0.028	-43.011	0	-1.259	-1.149	0.3	0.28	0.32
SUBDISCIPLINE[T.Nursing]	-1.2469	0.033	-37.548	0	-1.312	-1.182	0.29	0.27	0.31
SUBDISCIPLINE[T.Nutrition & Dietetic]	-1.3448	0.033	-41.064	0	-1.409	-1.281	0.26	0.24	0.28
SUBDISCIPLINE[T.Obstetrics & Gynecology]	-3.9056	0.036	-108.84	0	-3.976	-3.835	0.02	0.02	0.02
SUBDISCIPLINE[T.Ophthalmology]	-1.0694	0.031	-34.806	0	-1.13	-1.009	0.34	0.32	0.36
SUBDISCIPLINE[T.Orthopedics]	-1.298	0.031	-42.39	0	-1.358	-1.238	0.27	0.26	0.29
SUBDISCIPLINE[T.Otorhinolaryngology]	-0.912	0.032	-28.393	0	-0.975	-0.849	0.4	0.38	0.43
SUBDISCIPLINE[T.Parasitology]	-2.969	0.04	-73.875	0	-3.048	-2.89	0.05	0.05	0.06
SUBDISCIPLINE[T.Pathology]	-2.1439	0.03	-70.807	0	-2.203	-2.085	0.12	0.11	0.12
SUBDISCIPLINE[T.Pediatrics]	-1.0785	0.03	-35.518	0	-1.138	-1.019	0.34	0.32	0.36
SUBDISCIPLINE[T.Pharmacology]	-2.2187	0.028	-78.092	0	-2.274	-2.163	0.11	0.1	0.11

SUBDISCIPLINE[T.Pharmacy]	-2.5641	0.046	-55.445	0	-2.655	-2.473	0.08	0.07	0.08
SUBDISCIPLINE[T.Physiology]	-1.5502	0.034	-46.171	0	-1.616	-1.484	0.21	0.2	0.23
SUBDISCIPLINE[T.Psychiatry]	-0.5844	0.032	-18.106	0	-0.648	-0.521	0.56	0.52	0.59
SUBDISCIPLINE[T.Public Health]	-1.0154	0.031	-32.991	0	-1.076	-0.955	0.36	0.34	0.38
SUBDISCIPLINE[T.Radiology & Nuclear Medicine]	-1.764	0.029	-60.09	0	-1.822	-1.706	0.17	0.16	0.18
SUBDISCIPLINE[T.Rehabilitation]	-0.3991	0.037	-10.819	0	-0.471	-0.327	0.67	0.62	0.72
SUBDISCIPLINE[T.Respiratory System]	-1.2449	0.032	-38.361	0	-1.308	-1.181	0.29	0.27	0.31
SUBDISCIPLINE[T.Social Studies of Medicine]	-1.7652	0.208	-8.482	0	-2.173	-1.357	0.17	0.11	0.26
SUBDISCIPLINE[T.Speech-Language Pathology and Audiology]	0.2348	0.059	3.973	0	0.119	0.351	1.26	1.13	1.42
SUBDISCIPLINE[T.Surgery]	-1.3175	0.029	-46.174	0	-1.373	-1.262	0.27	0.25	0.28
SUBDISCIPLINE[T.Tropical Medicine]	-1.9723	0.039	-50.481	0	-2.049	-1.896	0.14	0.13	0.15
SUBDISCIPLINE[T.Urology]	-1.031	0.031	-33.149	0	-1.092	-0.97	0.36	0.34	0.38
SUBDISCIPLINE[T.Veterinary Medicine]	-2.3008	0.031	-73.307	0	-2.362	-2.239	0.1	0.09	0.11
SUBDISCIPLINE[T.Virology]	-3.077	0.034	-89.54	0	-3.144	-3.01	0.05	0.04	0.05
YEAR	0.0092	0.001	11.573	0	0.008	0.011	1.01	1.01	1.01
np.log2(N_AUTHORS)	0.6053	0.003	224.524	0	0.6	0.611	1.83	1.82	1.84
F_MESH	-3.7915	0.032	-119.43	0	-3.854	-3.729	0.02	0.02	0.02
F_COUNTRY	0.9186	0.036	25.613	0	0.848	0.989	2.51	2.33	2.69

676

677 **Table S8.** Coefficients and odds ratios of the variables from the logistic regression, sex-related  
678 reporting = male (SR\_F).

Logit Regression Results

Dep. Variable:	SR_F	Log-Likelihood:	7.8E+05
Model:	Logit	LL-Null:	8.7E+05
Method:	MLE	LLR p-value:	0.0E+00
No. Observations:	1273687		
Df Residuals:	1273618		
Df Model:	68		
Pseudo R-sq.	0.1021		

	coef	std err	z	P> z	[0.025	0.975]	Odds Ratio	[0.025	0.975]
Intercept	-27.413	1.584	-17.301	0	-30.519	-24.308	0	0	0
Male-Female	0.0756	0.006	12.437	0	0.064	0.088	1.08	1.07	1.09
Female-Male	0.075	0.005	15.673	0	0.066	0.084	1.08	1.07	1.09
Female-Female	0.2507	0.006	40.256	0	0.238	0.263	1.28	1.27	1.3
CONTINENT[T.Africa]	0.7045	0.024	29.262	0	0.657	0.752	2.02	1.93	2.12
CONTINENT[T.Asia]	0.203	0.006	33.225	0	0.191	0.215	1.23	1.21	1.24
CONTINENT[T.Europe]	0.1443	0.005	30.176	0	0.135	0.154	1.16	1.14	1.17
CONTINENT[T.Oceania]	0.0481	0.012	4.06	0	0.025	0.071	1.05	1.03	1.07
CONTINENT[T.South America]	-0.0766	0.012	-6.565	0	-0.099	-0.054	0.93	0.91	0.95
SUBDISCIPLINE[T.Allergy]	-0.6237	0.048	-13.112	0	-0.717	-0.53	0.54	0.49	0.59
SUBDISCIPLINE[T.Anatomy & Morphology]	-0.9236	0.064	-14.362	0	-1.05	-0.798	0.4	0.35	0.45
SUBDISCIPLINE[T.Anesthesiology]	-0.5692	0.037	-15.558	0	-0.641	-0.497	0.57	0.53	0.61
SUBDISCIPLINE[T.Arthritis & Rheumatology]	-0.3369	0.031	-10.708	0	-0.399	-0.275	0.71	0.67	0.76
SUBDISCIPLINE[T.Biochemistry & Molecular Biology]	-2.3817	0.028	-86.142	0	-2.436	-2.327	0.09	0.09	0.1
SUBDISCIPLINE[T.Biomedical Engineering]	-2.0946	0.04	-52.537	0	-2.173	-2.016	0.12	0.11	0.13
SUBDISCIPLINE[T.Biomedical Social Sciences]	-0.2241	0.072	-3.106	0.002	-0.366	-0.083	0.8	0.69	0.92
SUBDISCIPLINE[T.Biophysics]	-2.0579	0.061	-33.832	0	-2.177	-1.939	0.13	0.11	0.14
SUBDISCIPLINE[T.Cancer]	-0.8121	0.026	-30.885	0	-0.864	-0.761	0.44	0.42	0.47
SUBDISCIPLINE[T.Cardiovascular System]	-0.3878	0.027	-14.384	0	-0.441	-0.335	0.68	0.64	0.72
SUBDISCIPLINE[T.Cellular Biology Cytology & Histology]	-2.2116	0.031	-70.664	0	-2.273	-2.15	0.11	0.1	0.12
SUBDISCIPLINE[T.Dentistry]	-0.5256	0.031	-16.983	0	-0.586	-0.465	0.59	0.56	0.63
SUBDISCIPLINE[T.Dermatology & Venereal Disease]	-0.7931	0.029	-26.938	0	-0.851	-0.735	0.45	0.43	0.48
SUBDISCIPLINE[T.Embryology]	-0.9849	0.057	-17.167	0	-1.097	-0.872	0.37	0.33	0.42
SUBDISCIPLINE[T.Endocrinology]	-0.6221	0.028	-22.003	0	-0.677	-0.567	0.54	0.51	0.57
SUBDISCIPLINE[T.Environmental & Occupational Health]	-0.3665	0.033	-11.235	0	-0.43	-0.303	0.69	0.65	0.74

SUBDISCIPLINE[T.Fertility]	0.3898	0.046	8.418	0	0.299	0.481	1.48	1.35	1.62
SUBDISCIPLINE[T.Gastroenterology]	-0.6297	0.028	-22.52	0	-0.685	-0.575	0.53	0.5	0.56
SUBDISCIPLINE[T.General & Internal Medicine]	-0.7269	0.026	-27.489	0	-0.779	-0.675	0.48	0.46	0.51
SUBDISCIPLINE[T.General Biomedical Research]	-1.6078	0.027	-59.998	0	-1.66	-1.555	0.2	0.19	0.21
SUBDISCIPLINE[T.Genetics & Heredity]	-1.5997	0.029	-54.664	0	-1.657	-1.542	0.2	0.19	0.21
SUBDISCIPLINE[T.Geriatrics]	-0.134	0.043	-3.111	0.002	-0.218	-0.05	0.87	0.8	0.95
SUBDISCIPLINE[T.Geriatrics & Gerontology]	-0.3447	0.044	-7.9	0	-0.43	-0.259	0.71	0.65	0.77
SUBDISCIPLINE[T.Health Policy & Services]	-0.4	0.033	-12.159	0	-0.465	-0.336	0.67	0.63	0.71
SUBDISCIPLINE[T.Hematology]	-1.0087	0.03	-34.147	0	-1.067	-0.951	0.36	0.34	0.39
SUBDISCIPLINE[T.Immunology]	-1.457	0.027	-54.198	0	-1.51	-1.404	0.23	0.22	0.25
SUBDISCIPLINE[T.Microbiology]	-2.1838	0.031	-69.835	0	-2.245	-2.122	0.11	0.11	0.12
SUBDISCIPLINE[T.Microscopy]	-2.8678	0.201	-14.241	0	-3.262	-2.473	0.06	0.04	0.08
SUBDISCIPLINE[T.Miscellaneous Biomedical Research]	-1.3945	0.036	-38.696	0	-1.465	-1.324	0.25	0.23	0.27
SUBDISCIPLINE[T.Miscellaneous Clinical Medicine]	-0.6972	0.031	-22.271	0	-0.759	-0.636	0.5	0.47	0.53
SUBDISCIPLINE[T.Nephrology]	-0.7263	0.034	-21.506	0	-0.792	-0.66	0.48	0.45	0.52
SUBDISCIPLINE[T.Neurology & Neurosurgery]	-0.7799	0.026	-29.673	0	-0.831	-0.728	0.46	0.44	0.48
SUBDISCIPLINE[T.Nursing]	-0.2365	0.033	-7.266	0	-0.3	-0.173	0.79	0.74	0.84
SUBDISCIPLINE[T.Nutrition & Dietetic]	-0.9651	0.031	-30.924	0	-1.026	-0.904	0.38	0.36	0.4
SUBDISCIPLINE[T.Obstetrics & Gynecology]	2.131	0.047	45.448	0	2.039	2.223	8.42	7.68	9.23
SUBDISCIPLINE[T.Ophthalmology]	-0.1717	0.029	-5.89	0	-0.229	-0.115	0.84	0.8	0.89
SUBDISCIPLINE[T.Orthopedics]	0.1291	0.029	4.448	0	0.072	0.186	1.14	1.07	1.2
SUBDISCIPLINE[T.Otorhinolaryngology]	-0.0083	0.03	-0.271	0.787	-0.068	0.051	0.99	0.93	1.05
SUBDISCIPLINE[T.Parasitology]	-2.0134	0.037	-53.702	0	-2.087	-1.94	0.13	0.12	0.14

SUBDISCIPLINE[T.Pathology]	-0.9398	0.029	-32.692	0	-0.996	-0.883	0.39	0.37	0.41
SUBDISCIPLINE[T.Pediatrics]	-0.2864	0.029	-9.824	0	-0.344	-0.229	0.75	0.71	0.8
SUBDISCIPLINE[T.Pharmacology]	-1.9866	0.027	-73.7	0	-2.039	-1.934	0.14	0.13	0.14
SUBDISCIPLINE[T.Pharmacy]	-1.9839	0.046	-43.079	0	-2.074	-1.894	0.14	0.13	0.15
SUBDISCIPLINE[T.Physiology]	-1.7484	0.033	-53.727	0	-1.812	-1.685	0.17	0.16	0.19
SUBDISCIPLINE[T.Psychiatry]	0.1543	0.032	4.895	0	0.092	0.216	1.17	1.1	1.24
SUBDISCIPLINE[T.Public Health]	-0.2396	0.03	-8.027	0	-0.298	-0.181	0.79	0.74	0.83
SUBDISCIPLINE[T.Radiology & Nuclear Medicine]	-0.4744	0.028	-17.029	0	-0.529	-0.42	0.62	0.59	0.66
SUBDISCIPLINE[T.Rehabilitation]	0.0298	0.035	0.845	0.398	-0.039	0.099	1.03	0.96	1.1
SUBDISCIPLINE[T.Respiratory System]	-0.4963	0.031	-16.066	0	-0.557	-0.436	0.61	0.57	0.65
SUBDISCIPLINE[T.Social Studies of Medicine]	-1.1108	0.206	-5.385	0	-1.515	-0.707	0.33	0.22	0.49
SUBDISCIPLINE[T.Speech-Language Pathology and Audiology]	-0.2962	0.055	-5.385	0	-0.404	-0.188	0.74	0.67	0.83
SUBDISCIPLINE[T.Surgery]	-0.1298	0.027	-4.823	0	-0.183	-0.077	0.88	0.83	0.93
SUBDISCIPLINE[T.Tropical Medicine]	-1.1218	0.038	-29.576	0	-1.196	-1.047	0.33	0.3	0.35
SUBDISCIPLINE[T.Urology]	-1.0611	0.029	-36.913	0	-1.117	-1.005	0.35	0.33	0.37
SUBDISCIPLINE[T.Veterinary Medicine]	-1.3731	0.03	-46.245	0	-1.431	-1.315	0.25	0.24	0.27
SUBDISCIPLINE[T.Virology]	-2.1383	0.032	-66.545	0	-2.201	-2.075	0.12	0.11	0.13
YEAR	0.0128	0.001	16.251	0	0.011	0.014	1.01	1.01	1.01
np.log2(N_AUTHORS)	0.572	0.003	214.342	0	0.567	0.577	1.77	1.76	1.78
F_MESH	2.838	0.032	89.218	0	2.776	2.9	17.08	16.05	18.17
F_COUNTRY	0.8258	0.036	23.24	0	0.756	0.895	2.28	2.13	2.45

679  
680



681 **Table S9.** Coefficients and odds ratios of the variables from the logistic regression, sex-related  
 682 reporting = male (SR\_B).

Logit Regression Results

Dep. Variable:	SR_B	Log-Likelihood:	7.8E+05
Model:	Logit	LL-Null:	8.7E+05
Method:	MLE	LLR p-value:	0.0E+00
No. Observations:	1273687		
Df Residuals:	1273618		
Df Model:	68		
Pseudo R-sq.	0.1065		

	coef	std err	z	P> z	[0.025	0.975]	Odds Ratio	[0.025	0.975]
Intercept	-49.14	1.593	-30.84	0	-52.263	-46.017	0	0	0
Male-Female	0.0535	0.006	8.72	0	0.041	0.066	1.05	1.04	1.07
Female-Male	0.0348	0.005	7.222	0	0.025	0.044	1.04	1.03	1.04
Female-Female	0.0981	0.006	15.815	0	0.086	0.11	1.1	1.09	1.12
CONTINENT[T.Africa]	0.7188	0.023	31.282	0	0.674	0.764	2.05	1.96	2.15
CONTINENT[T.Asia]	0.1737	0.006	28.289	0	0.162	0.186	1.19	1.18	1.2
CONTINENT[T.Europe]	0.1796	0.005	37.377	0	0.17	0.189	1.2	1.19	1.21
CONTINENT[T.Oceania]	0.1291	0.012	10.934	0	0.106	0.152	1.14	1.11	1.16
CONTINENT[T.South America]	-0.0877	0.012	-7.513	0	-0.111	-0.065	0.92	0.89	0.94
SUBDISCIPLINE[T.Allergy]	-0.786	0.046	-17.148	0	-0.876	-0.696	0.46	0.42	0.5
SUBDISCIPLINE[T.Anatomy & Morphology]	-1.7987	0.066	-27.402	0	-1.927	-1.67	0.17	0.15	0.19
SUBDISCIPLINE[T.Anesthesiology]	-1.4526	0.035	-40.959	0	-1.522	-1.383	0.23	0.22	0.25
SUBDISCIPLINE[T.Arthritis & Rheumatology]	-0.9666	0.029	-32.823	0	-1.024	-0.909	0.38	0.36	0.4
SUBDISCIPLINE[T.Biochemistry & Molecular Biology]	-3.1046	0.027	-113.47	0	-3.158	-3.051	0.04	0.04	0.05
SUBDISCIPLINE[T.Biomedical Engineering]	-2.9319	0.043	-67.43	0	-3.017	-2.847	0.05	0.05	0.06
SUBDISCIPLINE[T.Biomedical Social Sciences]	-0.5493	0.068	-8.08	0	-0.683	-0.416	0.58	0.51	0.66
SUBDISCIPLINE[T.Biophysics]	-2.86	0.069	-41.186	0	-2.996	-2.724	0.06	0.05	0.07
SUBDISCIPLINE[T.Cancer]	-2.0529	0.025	-82.657	0	-2.102	-2.004	0.13	0.12	0.13
SUBDISCIPLINE[T.Cardiovascular System]	-1.0956	0.025	-43.176	0	-1.145	-1.046	0.33	0.32	0.35
SUBDISCIPLINE[T.Cellular Biology Cytology & Histology]	-3.1322	0.033	-94.071	0	-3.197	-3.067	0.04	0.04	0.05

SUBDISCIPLINE[T.Dentistry]	-1.1111	0.029	-37.811	0	-1.169	-1.053	0.33	0.31	0.35
SUBDISCIPLINE[T.Dermatology & Venereal Disease]	-1.4217	0.028	-50.552	0	-1.477	-1.367	0.24	0.23	0.25
SUBDISCIPLINE[T.Embryology]	-2.1862	0.061	-36.078	0	-2.305	-2.067	0.11	0.1	0.13
SUBDISCIPLINE[T.Endocrinology]	-1.2277	0.027	-46.214	0	-1.28	-1.176	0.29	0.28	0.31
SUBDISCIPLINE[T.Environmental & Occupational Health]	-0.7444	0.03	-24.47	0	-0.804	-0.685	0.48	0.45	0.5
SUBDISCIPLINE[T.Fertility]	-2.6492	0.044	-60.697	0	-2.735	-2.564	0.07	0.06	0.08
SUBDISCIPLINE[T.Gastroenterology]	-1.1335	0.026	-42.874	0	-1.185	-1.082	0.32	0.31	0.34
SUBDISCIPLINE[T.General & Internal Medicine]	-1.4042	0.025	-56.407	0	-1.453	-1.355	0.25	0.23	0.26
SUBDISCIPLINE[T.General Biomedical Research]	-2.1979	0.026	-86.178	0	-2.248	-2.148	0.11	0.11	0.12
SUBDISCIPLINE[T.Genetics & Heredity]	-2.0701	0.028	-72.92	0	-2.126	-2.014	0.13	0.12	0.13
SUBDISCIPLINE[T.Geriatrics]	-0.7792	0.039	-19.888	0	-0.856	-0.702	0.46	0.42	0.5
SUBDISCIPLINE[T.Geriatrics & Gerontology]	-0.5191	0.042	-12.499	0	-0.601	-0.438	0.6	0.55	0.65
SUBDISCIPLINE[T.Health Policy & Services]	-0.6884	0.031	-22.162	0	-0.749	-0.628	0.5	0.47	0.53
SUBDISCIPLINE[T.Hematology]	-1.4076	0.028	-50.072	0	-1.463	-1.353	0.24	0.23	0.26
SUBDISCIPLINE[T.Immunology]	-1.9646	0.026	-76.805	0	-2.015	-1.914	0.14	0.13	0.15
SUBDISCIPLINE[T.Microbiology]	-2.7113	0.032	-85.065	0	-2.774	-2.649	0.07	0.06	0.07
SUBDISCIPLINE[T.Microscopy]	-4.0137	0.313	-12.806	0	-4.628	-3.399	0.02	0.01	0.03
SUBDISCIPLINE[T.Miscellaneous Biomedical Research]	-2.2539	0.037	-60.779	0	-2.327	-2.181	0.1	0.1	0.11
SUBDISCIPLINE[T.Miscellaneous Clinical Medicine]	-1.197	0.03	-40.063	0	-1.256	-1.138	0.3	0.28	0.32
SUBDISCIPLINE[T.Nephrology]	-1.1619	0.032	-35.843	0	-1.225	-1.098	0.31	0.29	0.33
SUBDISCIPLINE[T.Neurology & Neurosurgery]	-1.2292	0.025	-49.727	0	-1.278	-1.181	0.29	0.28	0.31
SUBDISCIPLINE[T.Nursing]	-0.7669	0.03	-25.207	0	-0.827	-0.707	0.46	0.44	0.49

SUBDISCIPLINE[T.Nutrition & Dietetic]	-1.4068	0.03	-47.029	0	-1.465	-1.348	0.24	0.23	0.26
SUBDISCIPLINE[T.Obstetrics & Gynecology]	-3.3568	0.034	-98.636	0	-3.423	-3.29	0.03	0.03	0.04
SUBDISCIPLINE[T.Ophthalmology]	-0.6838	0.027	-24.935	0	-0.738	-0.63	0.5	0.48	0.53
SUBDISCIPLINE[T.Orthopedics]	-0.7561	0.027	-27.684	0	-0.81	-0.703	0.47	0.44	0.5
SUBDISCIPLINE[T.Otorhinolaryngology]	-0.5412	0.029	-18.877	0	-0.597	-0.485	0.58	0.55	0.62
SUBDISCIPLINE[T.Parasitology]	-2.6412	0.04	-65.348	0	-2.72	-2.562	0.07	0.07	0.08
SUBDISCIPLINE[T.Pathology]	-2.0723	0.028	-74.603	0	-2.127	-2.018	0.13	0.12	0.13
SUBDISCIPLINE[T.Pediatrics]	-0.8341	0.027	-30.683	0	-0.887	-0.781	0.43	0.41	0.46
SUBDISCIPLINE[T.Pharmacology]	-2.5816	0.026	-99.451	0	-2.633	-2.531	0.08	0.07	0.08
SUBDISCIPLINE[T.Pharmacy]	-2.5989	0.05	-51.747	0	-2.697	-2.5	0.07	0.07	0.08
SUBDISCIPLINE[T.Physiology]	-2.3315	0.033	-71.466	0	-2.395	-2.268	0.1	0.09	0.1
SUBDISCIPLINE[T.Psychiatry]	-0.3238	0.029	-11.238	0	-0.38	-0.267	0.72	0.68	0.77
SUBDISCIPLINE[T.Public Health]	-0.718	0.028	-25.923	0	-0.772	-0.664	0.49	0.46	0.51
SUBDISCIPLINE[T.Radiology & Nuclear Medicine]	-1.4577	0.026	-55.536	0	-1.509	-1.406	0.23	0.22	0.25
SUBDISCIPLINE[T.Rehabilitation]	-0.1489	0.033	-4.492	0	-0.214	-0.084	0.86	0.81	0.92
SUBDISCIPLINE[T.Respiratory System]	-0.8489	0.029	-28.983	0	-0.906	-0.791	0.43	0.4	0.45
SUBDISCIPLINE[T.Social Studies of Medicine]	-1.4679	0.205	-7.144	0	-1.871	-1.065	0.23	0.15	0.34
SUBDISCIPLINE[T.Speech-Language Pathology and Audiology]	0.1717	0.053	3.23	0.001	0.068	0.276	1.19	1.07	1.32
SUBDISCIPLINE[T.Surgery]	-1.0708	0.025	-42.414	0	-1.12	-1.021	0.34	0.33	0.36
SUBDISCIPLINE[T.Tropical Medicine]	-1.5716	0.037	-42.326	0	-1.644	-1.499	0.21	0.19	0.22
SUBDISCIPLINE[T.Urology]	-1.8931	0.028	-68.562	0	-1.947	-1.839	0.15	0.14	0.16
SUBDISCIPLINE[T.Veterinary Medicine]	-2.1095	0.029	-71.862	0	-2.167	-2.052	0.12	0.11	0.13

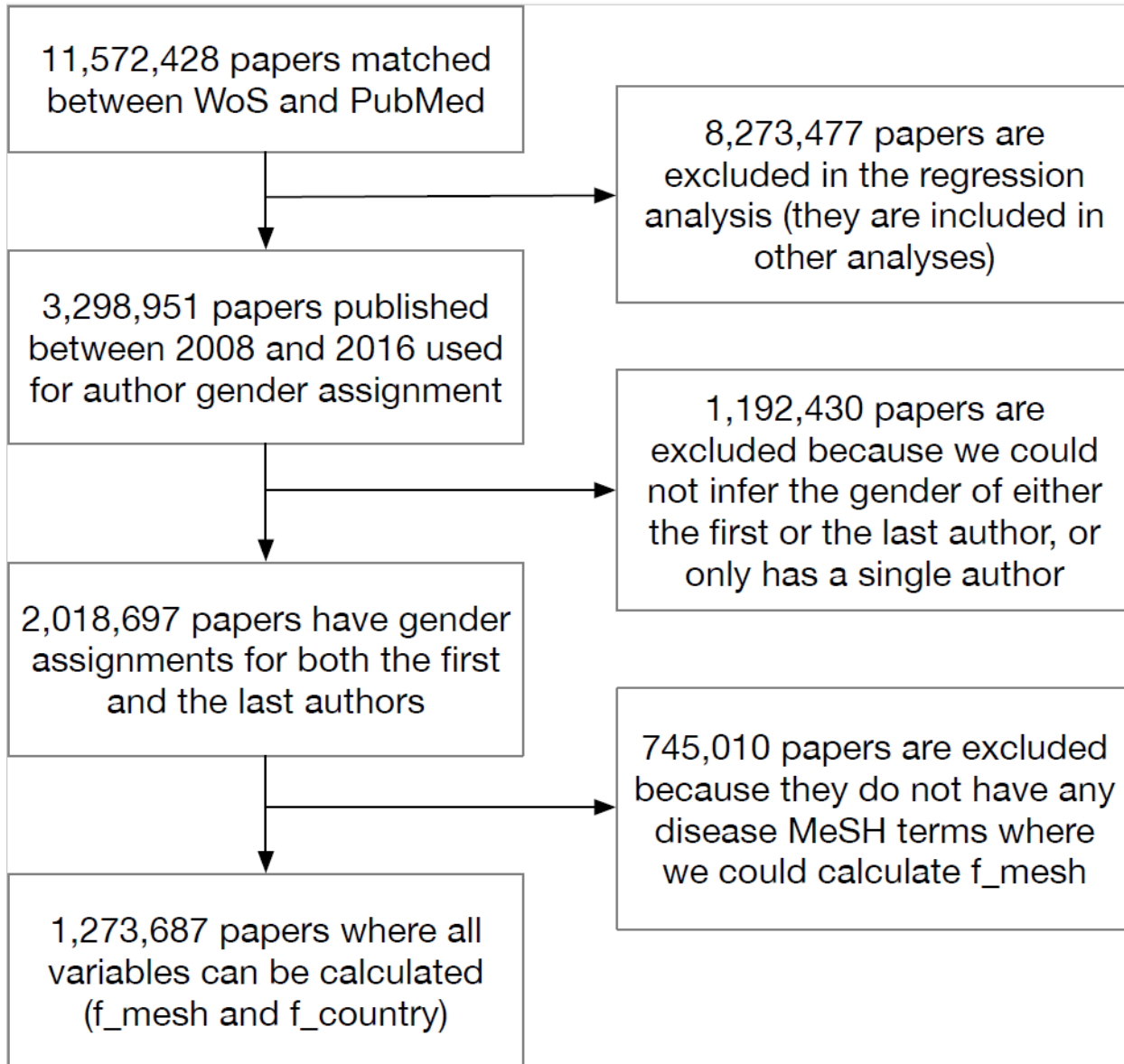
SUBDISCIPLINE[T.Virology]	-2.6294	0.033	-80.28	0	-2.694	-2.565	0.07	0.07	0.08
YEAR	0.0243	0.001	30.731	0	0.023	0.026	1.02	1.02	1.03
np.log2(N_AUTHORS)	0.5982	0.003	219.658	0	0.593	0.604	1.82	1.81	1.83
F_MESH	-1.6906	0.031	-53.896	0	-1.752	-1.629	0.18	0.17	0.2
F_COUNTRY	1.0789	0.035	30.55	0	1.01	1.148	2.94	2.75	3.15

---

683

684

685 **Figure S1. Creation of the dataset for the regression analysis**  
686



687  
688