Gender Diverse Teams Produce More Innovative and Influential Ideas in Medical Research

Yang Yang, PhD[1,2], Teresa Woodruff, PhD[3], Yuan Tian, PhD[4], Benjamin F. Jones, PhD[2,5,6], Brian Uzzi, PhD[2,5,7*]

Syracuse University[1]
Northwestern Institute on Complex Systems[2]
Michigan State University[3]
New York University[4]
Kellogg School of Management, Northwestern University[5]
National Bureau of Economic Research[6]
McCormick School of Engineering, Northwestern University[7]

*corresponding author: Brian Uzzi, email: uzzi@northwestern.edu

Words Count: 2,208

Key Points

Question: Is a research team's gender composition related to its innovativeness and impact?

Findings: Gender diverse teams are significantly more innovative and impactful than same gender teams. The advantages of gender diverse teams generalize to medical science subfields, hold whether the gender of the diverse team's leader is a man or women, and increase as the team's gender mix becomes more balanced. Nevertheless, gender diverse teams are still significantly underrepresented in medical science, potentially inhibiting the advancement of science and women in science.

Meaning: Gender diverse teams play an unrecognized key role in medical science's innovative and high impact research.

Abstract

Importance: This work addresses questions about how concurrent increases in teamwork[1] and women in science[2,3] have changed how the gender composition of a team positively coincides with its research innovativeness, citation impact, and potential for recognizing women's new contributions to science[4-11].

Objectives: Statistically examine the link between gender diverse research teams and team performance relative to same gender teams.

Design: Systemic study of team science using a massive, longitudinal scientometric data covering all 6.6 million publications from 15,033 medical science journals and published by 3.2 million female and 4.4 million male authors, 2000-2019.

Setting: International sample of 7.6 million medical researchers.

Participants: All authors of academic medical research published from 2000-2019 in 15,033 medical journals worldwide as curated in Microsoft Academic Graph (MAG)[12].

Design: We regressed each publication's innovativeness and normalized citation impact on the gender composition of the team using fixed-effects, matching, and null models to control for confounds due to institutional, journal, time, and team and individual performance variables.

Main outcomes and measures: Research performance is measured in terms of innovativeness and impact and the representativeness of gender diverse teams.

Results: Gender diverse teams publish papers that are up to 7% more innovative than same gender teams, controlling for confounds due to the authors' past performance, prestige, subfield, and journal. Similarly, gender diverse teams can be twice as likely to publish hit papers (top 5% of citations) relative to the base rate and 16.7% more likely than same-gender teams. The results strongly generalize to the subfields of medical science, whether the diverse team's leader is a woman or a man, and further strengthen as gender diverse teams become more gender balanced.

Conclusions: Gender diverse teams outperform same gender teams in the creation of innovative and high impact papers, an advantage that generalizes to medical science subfields, holds whether the gender of the diverse team's leader is a man or women, and increases as the team's gender mix becomes more balanced. Despite their advantages, gender diverse teams are significantly underrepresented in medical science, potentially constraining scientific advances and under-utilizing human capital.

Introduction

In the past 20 years, medicine has undergone two transformations that may potentially be remaking medical research outcomes. One change involves a shift from individual to team science[1]. Rising teamwork levels are broadly documented in science. Teams produce more highly cited research and influence stratification[13], yet the implications of team science for medical research has yet to be fully investigated[14,15]. The second change is the increased participation of women in medical science[3,4,16]. Perhaps more than any other branch of science, the last decade has seen women's participation rates exceed the participation rates of men in graduate and post-doctoral training in 60%-to-40% and 54%-to-46% ratios, respectively[17].

These systemic changes create a need to study how a research team's gender composition[18,19] is related to research performance[2,20] and acknowledgement of credit[11,21-24]. Is a medical science team's performance correlated with gender diversity? Are the implied gains of gender diverse teams measurable? Does team performance vary with a leader's gender?

New experiments suggest that gender diverse teams are potentially consequential to the innovativeness and impact of research because they may combine the benefits of teams and diversity. Lab experiments indicate that mixed-gender teams exchange more diverse information than same gender teams and can do better at general problem-solving tasks than all-male or all-female teams with equivalent IQ levels[8]. Yet, when gender stereotypes and inequities exist in professional settings like medicine, same gender relationships are preferred[25-27], reducing workgroup diversity and equitable representation[28]. These theoretical and practical repercussions merit new empirical analysis of the incidence and impact of gender diverse teams on medical science research and practice.

Methods

Sample

We conduct the first large-scale systemic investigation of the performance of gender diverse research teams in the medical sciences. Our original field-wide dataset has over 6.6 million research publications by 3,225,279 female and 4,386,020 male scientists in 15,033 medical science journals from 2000-2019 as recorded in Microsoft Academic Graph[12]. Microsoft Academic Graph (MAG) is a scientific publication database that records journal article's bibliographic information (title, journal, journal field, volume, issue, page, publication date), authorship (name), author affiliations (name, webpage, and wikipage), and citation links to other papers in the database.

Variables

Our main independent variable is team gender diversity, and our two outcome variables are research innovativeness and research impact. All variables were constructed with MAG data.

Team Gender Diversity: To measure the gender composition in a scientific team, we use a binary variable "mixed-gender". A mixed-gender team has both men and women. Otherwise, it is a same-gender team. We also use a continuous variable to evaluate the gender composition of a scientific team in finer detail[26] that takes the form

$$g_i = -p_f log_2(p_f) - (1 - p_f) log_2(1 - p_f) \quad (1)$$

where p_f indicates the portion of female authors in a team i. The value of g_i ranges from 0 to 1. When the value of g_i is low, either women or men are majority of a team. When $g_i = 0$, the team is either an all-women team or an all-men team. By contrast, when the value of g_i is high, women and men have roughly equivalent presence in the team. When $g_i = 1$, the team has 50% women and 50% men. A given author's gender was estimated using an accepted and validated algorithm[29]. Further supplementary validity tests demonstrated that algorithm's results are robust to measurement error (see **SM S1.2** and **Table S3** & **S4**).

Innovativeness of Scientific Papers: We use a popular and validated innovation measure that uses a paper's listed references as an indicator of its mixture of knowledge[30,31]. The measure considers papers with statistically atypical combinations of references to be innovative because they create new combinations of knowledge that have not been joined, or rarely joined, in previous research. By contrast, papers that have statistically common combinations of references represent conventional, familiar groupings of knowledge. To compute a paper's innovativeness in terms of novelty vs conventionality, we use a z-score. The z-score is computed from the observed frequency of journal pairs that appear within a paper's reference list and the expected distribution of journal pairings created by randomized citation networks. When a z-score is less than zero, the combination of prior work is considered innovative (i.e., a novel pairing of ideas) and when the z-score is above zero, the combination of prior work is considered conventional (i.e., a common pairing of ideas). A paper is defined as innovative if its novelty z-score is smaller than zero (see **SM S1.2.2** for detail).

Impact of Scientific Papers: The MAG database tracks a paper's annual citations. We define high impact papers as those in the top 5% of citation for papers published in the same year. To provide a fair and comparable measure, we normalize paper i's final citations by the publication year (see **SM S1.2.3** for detail).

Analysis

To quantify the link between team diversity, leadership, and performance, we used fixed effects regressions. The regressions control for confounds due to the authorship team's size, leadership, and institutional prestige rank, year, journal quality, prior citation impact, average team age, and individual fixed effects[1,13,30,32,33] and were run for the full data. To control for subfield differences, we also run a separate regression for 45 medical subfields (see **SM S2** for detail). Alternative measurements and null models confirm the results (see **SM S1** and **S3** for detail).

Results

Figure 1a shows team size trends in medical science. Large teams (6+ authors) are now the modal authorship form. Today almost 50% of all papers are by large teams, and the shares of papers at all other team sizes are concomitantly declining[1,34]. The share of publications by gender diverse teams is also changing[35]. We measure gender diversity two ways. First, "mixed-gender" is a binary variable equal to 0 if the team is all the same gender and 1 if the team has both female and male authors. Second, "gender diversity" is a continuous variable that rises from 0 when the team is all the same gender to 1 when the team includes equal numbers of women and men. **Supplemental Materials S1.2.4** describes the measures in detail. **Figures 1b** and **1d** indicate

that the share of papers by mixed-gender and gender-balanced teams respectively is growing with time.

However, **Figure 1c** and **1e** show that gender diverse are significantly underrepresented (means with 95% CI shown, all p-values < 0.01). The expected level of gender diverse teams was computed using a null model. The model holds constant the yearly observed share of female authors in medical science (i.e., ~45% in 2019), the coauthorship network, team size, and number of publications and randomly interchanges on a yearly basis female and male authors who are matched on having the same first year of publication, total publications, and country into simulated groups (see **SM S3.1** for details).

Gender Diverse Teams and Innovation and Impact

Despite their underrepresentation, gender diverse teams are more innovative and impactful than same gender teams. Innovativeness is measured by whether a paper combines past knowledge in a novel way relative to past combinations and is updated yearly to account for the evolving corpus of research and how it has been combined[30]. Overall, 44% of papers in our data are defined as innovative. High impact papers are in the top 5% of citation for papers normalized by yearly average citations (see **SM S1.2**).

Figures 2a and **2b** show the predicted margins of a paper's novelty and citation impact conditional on the authorship team's gender diversity and team size, while controlling for institutional prestige, the authors' prior citation impact (the mean for the team members and the prior citation impact of the first author and the last author separately), the authors' career age (the mean for the team members and that of the first author and last author separately), the gender of the first and last authors, journal-year fixed effects, and individual fixed effects (see **SM S2** for detail).

Figure 2a demonstrates that mixed-gender teams are significantly more likely to publish more novel papers than same-gender teams (two sample t-test, p-value < 0.001). For example, large (6+) mixed-gender team are 9.1% more likely to publish a novel article than the base rate [(0.48-0.44)/0.44]. Given that novelty positively correlates with team size[1,30,36], the substantial added explanatory power of mixed- vs. same-gender teams seen in **Figure 2** is striking. Proportionally, the increase in novelty for mixed-gender teams of 6+ relative to same-gender teams is equivalent to the increase in novelty obtained by doubling a same-gender team size from 2 to 4.

Figure 2b shows that mixed-gender teams are significantly more likely to publish a citation hit than same gender teams. Comparing increases in publishing a hit to the baseline rate, large (6+) mixed-gender teams are 16.7% [(10.5%-9.0%)/9.0%) =16.7%] more likely to publish a hit paper than same gender teams of equal size (two sample t-test, p-value < 0.001).

The results replicate when we examine the link between the continuous measure of team gender balance and performance. We find that as team composition moves from same gender to gender balanced, the positive impact on a paper's innovativeness and impact significantly increases irrespective of team size. **SM Tables S1** and **S2** present regression details and exact significance levels while controlling for confounds.

Lastly, we examined the generalizability of the link between team gender diversity and performance. First, we tested whether the gender of the team's first or last author matters[37]. We found that gender diverse teams led by men or women alike produce significantly more novel and higher impact research than same gender teams (two sample t-test, $p < 0.001$, see **SM S4**). Second, we tested whether the benefits of gender diverse teams generalize across medical research subfields. We grouped papers by their MAG designated primary subfield and ran a separate regression for each subfield using the same specification as described in **Figure 2**.

Figure 3 shows that the findings strongly generalize across subfields. The y-axis shows the regression coefficient value with 95% CIs when innovativeness (**Figure 3a**) and citation impact (**Figure 3b**) are regressed on the team's "mixed-gender" variable for 45 separate subfields (x-axis). **Figure 3a** and **3b** demonstrate that the team's gender diversity significantly and positively predicts a team's novelty and impact for a significant majority of subfields (binomial test $p < 0.03$ for mixed-gender and $p < 0.001$ for gender balance) with the smallest subfields exhibiting a noisy relationship. The continuous measure of gender balance replicates the findings of the mixed-gender measure (see **SM S2.6**).

Discussion

This research presents striking performance advantages of gender diverse teams relative to same-gender teams that generalizes broadly across subfields and whether the diverse team's leader is a woman or a man. The large body of work documenting the differences in the careers and advancement of women scientists has raised awareness of gender inequities that inhibit science, reduce workplace fairness, and require new policy[2-4]. This work reveals new, team-level patterns of gender relations in science. Teams that combine the efforts and talents of men and women scientists do better than either all men or all women teams. Gender diverse teams publish papers that are up to 7% more innovative and 16.7% more likely to be citation hits than same-gender teams, controlling for confounds due to the authors' past performance, prestige, subfield, and journal. These results cannot be explained by the increased frequency of teamwork, team size, or the surge in women's participation in medical science. Thus, this work provides a new perspective on the potentially transformative benefits of gender diversity in science.

While the research finds that gender diverse research teams have robust performance advantages over teams of all women or all men, our analysis also reveals that gender diverse teams remain significantly underrepresented in medical science. This suggests that medical science may have the potential to speed breakthroughs by breaking down barriers to the formation of gender diverse teams.

Relatedly, the underrepresentation of gender-diverse teams highlights earlier work that has documented the inequities women experience in relation to accurate perceptions of credit[11]. The reasons for underrepresentation may relate to gender inequities in grants[5], prizes[4], leaky pipelines[2], and credit allocation in teams[3,24] that have been identified elsewhere but bear further study and evaluation from the lens of team gender diversity. Bias in team formation may be related to these other challenges and point to mechanisms and options for instituting practices that can advance gender balance[11,38]. For example, adopting practices for listing each author's contribution in publications can further transparency, accountability, and fairness[11] that can otherwise inhibit team formation.[38] In lab experiments, it has been shown that misperceptions

can be ameliorated by providing feedback about individual team member performance[23]. Similarly, research examining causal tests of gender diversity in teams and mechanisms that inhibit gender-diverse team assembly is welcomed and could clarify actionable practices and support new policy[34].

Conclusions

Gender diverse teams are significantly more likely to publish research that innovatively combines existing ideas in new ways and is more influential than research by equivalent same gender teams. The advantages of mixed-gender teams generalize to medical science subfields, hold whether the gender of the diverse team's leader is a man or women, and increase as the team's gender mix becomes more balanced. Nevertheless, gender diverse teams are significantly underrepresented, potentially constraining scientific advances. More generally, this work expands science of science studies from analyses of gender differences to analyses of gender complementarities, recognizing distinctive advantages in the scientific outcomes from gender-diverse teams.

References

1. Wuchty S, Jones BF, Uzzi B. The increasing dominance of teams in production of knowledge. Science 2007;316:1036-9.
2. Etzkowitz H, Kemelgor C, Neuschatz M, Uzzi B, Alonzo J. The paradox of critical mass for women in science. Science 1994;266:51-3.
3. Etzkowitz H, Kemelgor C, Uzzi B. Athena Unbound: The Advancement of Women in Science and Technology. Cambridge: Cambridge University Press; 2000.
4. Ma Y, Oliveira DFM, Woodruff TK, Uzzi B. Women who win prizes get less money and prestige. Nature Publishing Group; 2019.
5. Oliveira DFM, Ma Y, Woodruff TK, Uzzi B. Comparison of National Institutes of Health Grant Amounts to First-Time Male and Female Principal Investigators. JAMA 2019;321:898-900.
6. Ma Y, Mukherjee S, Uzzi B. Mentorship and protégé success in STEM fields. Proceedings of the National Academy of Sciences 2020.
7. Woolley AW, Gerbasi ME, Chabris CF, Kosslyn SM, Hackman JR. Bringing in the experts - How team composition and collaborative planning jointly shape analytic effectiveness. Small Group Research 2008;39:352-71.
8. Woolley AW, Chabris CF, Pentland A, Hashmi N, Malone TW. Evidence for a collective intelligence factor in the performance of groups. Science 2010;330:686-8.
9. Woolley AW, Bear JB, Chang JW, DeCostanza AH. The effects of team strategic orientation on team process and information search. Organizational Behavior and Human Decision Processes 2013;122:114–26.
10. Nielsen MW, Alegria S, Börjeson L, et al. Opinion: Gender diversity leads to better science. Proceedings of the National Academy of Sciences 2017;114:1740-2.
11. Sugimoto C. Is science built on the shoulders of women? a study of gender differences in contributorship. Acad Med 2016;91.
12. Wang K, Shen Z, Huang C-Y, et al. A Review of Microsoft Academic Services for Science of Science Studies. Frontiers in Big Data 2019;2:45.
13. Jones BF, Wuchty S, Uzzi B. Multi-University Research Teams: Shifting Impact, Geography, and Stratification in Science. Science 2008;322:1259-62.
14. Fontanarosa P, Bauchner H, Flanagin A. Authorship and team science. Jama 2017;318:2433-7.
15. Falk-Krzesinski HJ, Börner K, Contractor N, et al. Advancing the science of team science. Clinical and translational science 2010;3:263-6.
16. Ataman LM, Ma Y, Duncan FE, Uzzi B, Woodruff TK. Quantifying the growth of oncofertility. Biology of reproduction 2018;99:263-5.
17. U.S. Medical School Faculty Trends: Percentages. 2020. at https://www.aamc.org/data-reports/faculty-institutions/interactive-data/us-medical-school-faculty-trends-percentages.)
18. Long JS. Measures of sex differences in scientific productivity. Social Forces 1992;71:159-78.
19. Bear JB, Woolley AW. The role of gender in team collaboration and performance. Interdisciplinary science reviews 2011;36:146-53.
20. Campbell LG, Mehtani S, Dozier ME, Rinehart J. Gender-heterogeneous working groups produce higher quality science. PloS one 2013;8:e79147.
21. Jagsi RG, Elizabeth A; Worobey, Cynthia Coooper; Henault, Lori E; Chang, Yuchiao; Starr, Rebecca; Tarbell, Nancy J; Hylek, Elaine M. The gender gap in authorship of academic medical literaturea 35-year perspective. New England Journal of Medicine 2006;355:281-7.

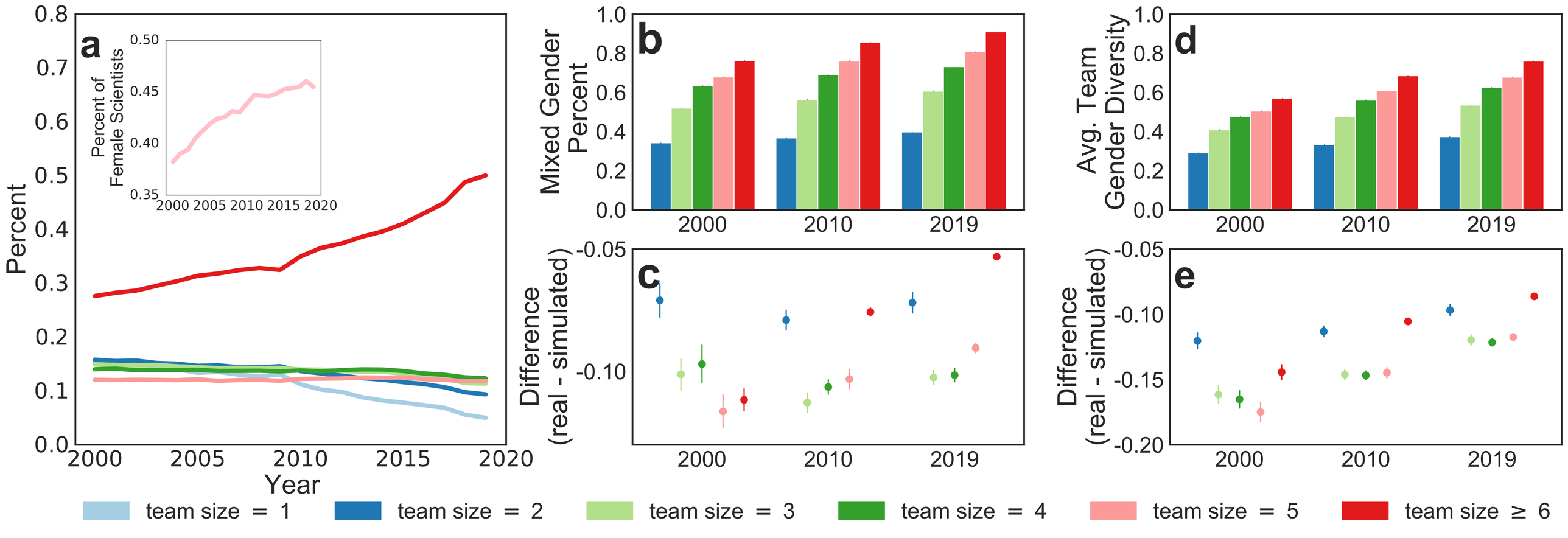
22. Patton EWG, Kent A; Jones, Rochelle D; Stewart, Abigail; Ubel, Peter A; Jagsi, Reshma. Differences in mentor-mentee sponsorship in male vs female recipients of national institutes of health grants. JAMA Internal Medicine 2017;177:580-2.
23. Heilman ME, Haynes MC. No credit where credit is due: attributional rationalization of women's success in male-female teams. Journal of applied Psychology 2005;90:905.
24. Sarsons H. Recognition for group work: Gender differences in academia. American Economic Review 2017;107:141-45.
25. Gorman EH. Gender stereotypes, same-gender preferences, and organizational variation in the hiring of women: Evidence from law firms. American Sociological Review 2005;70:702-28.
26. Yang Y, Chawla NV, Uzzi B. A network's gender composition and communication pattern predict women's leadership success. Proceedings of the National Academy of Sciences 2019;116:2033-8.
27. DeCastro R, Sambuco D, Ubel PA, Stewart A, Jagsi R. Mentor networks in academic medicine: moving beyond a dyadic conception of mentoring for junior faculty researchers. Academic medicine: journal of the Association of American Medical Colleges 2013;88:488.
28. Rivera LA. Hiring as cultural matching: The case of elite professional service firms. American Sociological Review 2012;77:999-1022.
29. Namsor. Available from: https://github.com/namsor/namsor-api. Retrieved October 30, 2020; 2020.
30. Uzzi B, Mukherjee S, Stringer M, Jones B. Atypical combinations and scientific impact. Science 2013;342:468-72.
31. Leahey E, Beckman CM, Stanko TL. Prominent but less productive: The impact of interdisciplinarity on scientists' research. Administrative Science Quarterly 2017;62:105-39.
32. Börner K, Contractor N, Falk-Krzesinski HJ, et al. A Multi-Level Systems Perspective for the Science of Team Science. Science Translational Medicine 2010;2:49cm24.
33. Jones B, Reedy EJ, Weinberg BA. Age and Scientific Genius. In: Simonton DK, ed. The Wiley Handbook of Genius: Wiley-Blackwell; 2014.
34. Fortunato S, Bergstrom CT, Börner K, et al. Science of science. Science 2018;359.
35. Ma Y, Uzzi B. Scientific prize network predicts who pushes the boundaries of science. Proceedings of the National Academy of Sciences 2018;115:12608-15.
36. Boudreau KJ, Guinan EC, Lakhani KR, Riedl C. Looking across and looking beyond the knowledge frontier: Intellectual distance, novelty, and resource allocation in science. Management Science 2016;62:2765-83.
37. Chowdhary M, Chowdhary A, Royce TJ, et al. Women's Representation in Leadership Positions in Academic Medical Oncology, Radiation Oncology, and Surgical Oncology Programs. JAMA Network Open 2020;3:e200708-e.
38. Ahmadpoor M, Jones BF. Decoding team and individual impact in science and invention. Proceedings of the National Academy of Sciences 2019;116:13885-90.

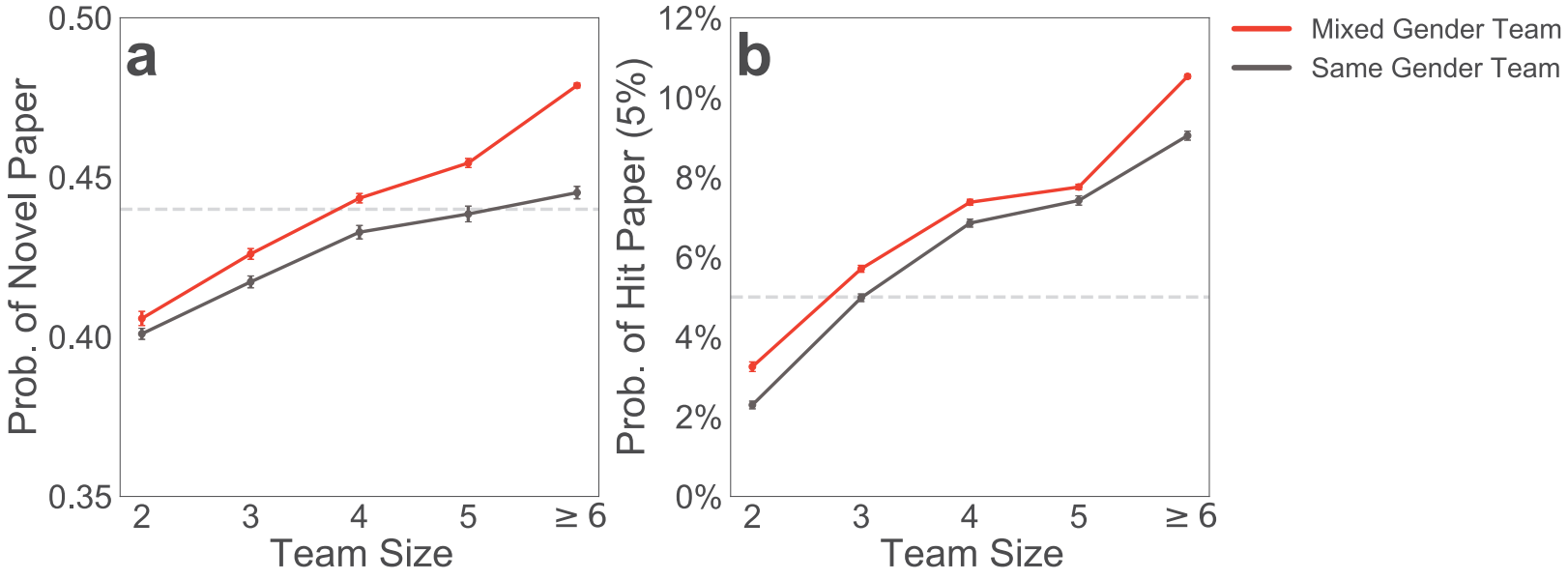
Figures

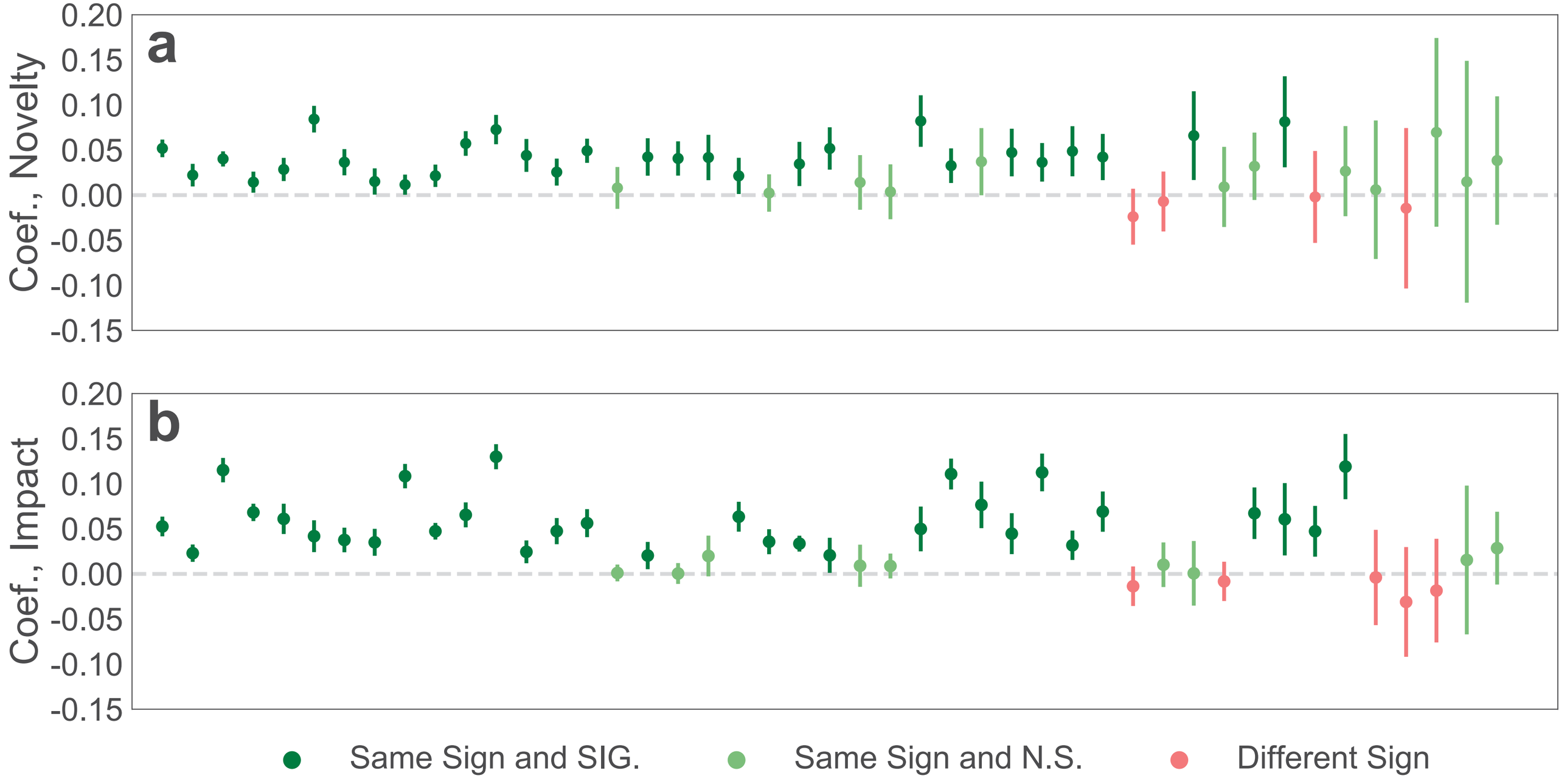
Figure 1: Gender Diverse Teams Remain Underrepresented in Medical Science. Figure 1a plots the share of publications (y-axis) by team size and year from 2000-2019 (x-axis) and figure 1a inset demonstrates the sharp upward trend of women's participation in medical science over the same time period. Large teams have replaced small teams. In 2000, solo and two-person teams accounted for 15% and 16% of publications, respectively, but by 2015 their shares dropped to 8% and 12% while the largest teams increased their market share dominance from 25% to 46%. Figure 1b and 1d shows that the share of publications from gender diverse teams steadily increased with time. Nonetheless, Figure 1c and 1e indicate that mixed-gender and gender balanced teams are significantly underrepresented in medical science by up to 17% depending on the team size.

Figure 2: Gender Diverse Teams Produce the Most Innovative and Highly Cited Research Publications. Mixed-gender teams are more likely to produce innovative papers than same-gender teams at all team sizes; for teams of size 4 or more, mixed-gender teams are also more likely to produce a novel paper than the base rate (dashed line). Mixed-gender teams of size 6+ (about 1/2 of all 2019 publications) are 9.1% more likely to publish novel work than the base rate. Figure 2b shows that the same performance patterns are largely mirrored for whether a paper is a citation hit (top 5% of papers). Large (6+) mixed-gendered teams can be twice as likely to publish hit papers than the base rate and 16.7% more likely than large (6+) same-gender teams [(10.5%-9.0%)/9.0%) = 16.7%]. The results are reproduced when team gender diversity is measured in terms of Shannon Entropy and shows that as gender balance of a team increases, innovativeness and impact significantly increase (see **SM Tables S1** and **S2**).

Figure 3: The Benefits of Gender Diverse Teams Generalize Across Medical Science Subfields. The plots demonstrate that the performance benefits of team gender diversity generalize across medical subfields. Figure a and b show the regression coefficients of 45 separate regressions. Each regression estimates the relationship between either novelty (figure a) or citation impact (figure b) and the team's gender diversity for the papers in one of 45 separate medical subfields. The regressions control for confounds due to the authors' experience, prestige, subfield, individual success, journal of publication and the gender of the team's leadership as was done in Figure 2, which pooled the data for subfields in medical science. Subfields are arranged from left to right according to the number of papers in the subfield (largest to smallest). The larger subfields show positive and statistically significant relationships with novelty and impact while the relationships become noisier with smaller fields with team gender diversity being significantly predictive of team performance for the majority of subfields, a statistical regularity that is not explained by chance (binomial test $p < 0.03$ for mixed-gender and $p < 0.001$ for gender balance).


a
Percent of Female Scientists
0.50
0.45
0.40
0.35
2000 2005 2010 2015 2020
Percent
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
2000
2005
2010
2015
2020
Year
b
Mixed Gender Percent
1.0
0.8
0.6
0.4
0.2
0.0
2000
2010
2019
c
Difference (real - simulated)
-0.05
-0.10
2000
2010
2019
d
Avg. Team Gender Diversity
1.0
0.8
0.6
0.4
0.2
0.0
2000
2010
2019
e
Difference (real - simulated)
-0.05
-0.10
-0.15
-0.20
2000
2010
2019
team size = 1
team size = 2
team size = 3
team size = 4
team size = 5
team size ≥ 6


a
Prob. of Novel Paper
0.50
0.45
0.40
0.35
2
3
4
5
≥ 6
Team Size
b
Prob. of Hit Paper (5%)
12%
10%
8%
6%
4%
2%
0%
2
3
4
5
≥ 6
Team Size
Mixed Gender Team
Same Gender Team


a
b
Coef., Novelty
Coef., Impact
0.20
0.15
0.10
0.05
0.00
-0.05
-0.10
-0.15
Same Sign and SIG.
Same Sign and N.S.
Different Sign