

Ethnicity and gender trends of UK authors in *The British Medical Journal* and the *Lancet* over the past two decades: a comprehensive longitudinal analysis



Salwa Abdalla,^{a,h} Moustafa Abdalla,^{b,c,h} Mohamed Saad,^d David Jones,^{b,e} Scott Podolsky,^{b,f} and Mohamed Abdalla^{a,g,*}

^aDepartment of Computer Science, University of Toronto, Toronto, Canada

^bHarvard Medical School, Boston, MA, USA

^cDepartment of Statistics, University of Oxford, Oxford, UK

^dUniversity of Bahrain & The Royal Academy, Manama, Bahrain

^eDepartment of the History of Science, Harvard University, Cambridge, MA, USA

^fCenter for the History of Medicine, Countway Medical Library, Harvard Medical School, Boston, MA, USA

^gInstitute for Better Health, Trillium Health Partners, Mississauga, Canada



Summary

Background While gender equity among academic authors has been extensively investigated, there is a significant gap in our understanding of racial/ethnic authorship trends, despite the recognition of barriers to authorship along both ethnic and gender lines. Leveraging the meta-data for all articles published in *The British Medical Journal* (*The BMJ*) and the *Lancet* and between 2002 and 2022 (inclusive), we explore demographic trends among UK academic medicine authors in two of the world's leading British medical journals.

Methods We systematically searched PubMed's MEDLINE for all articles published in *The BMJ* and *Lancet* between January 1st 2002 and December 31st 2022. Filtering for articles with a UK affiliation, we predicted gender using a publicly-validated name-to-gender dictionary, while data was analysed to explore and investigate ethnicity using the Consumer Data Research Centre's (CDRC) Ethnicity Estimator. Data was analysed to explore and investigate: (a) the proportion of female/male author publications, (b) the proportion of the various UK author ethnicity groups, and (c) the overlap/intersection between gender and ethnic identities among first and last authors. This comprehensive longitudinal analysis was conducted on 82,143 articles (51,209 from *The BMJ* and 30,934 from the *Lancet*) which represents >97% of all published articles between 2002 and 2022. As we sought to understand how academic authorship reflects the diversity of the UK population, we limited our analysis to first and last authors who had a UK affiliation and excluded "news" and "comments" pieces (16,736 articles for *The BMJ* and 4678 articles from the *Lancet*). The main outcome measures were the trends in first and last authorship demographics of academic medicine, focusing on the proportion of female/male authors, ethnicity and their intersectionality.

Findings Our findings show that, while women have made substantial headway towards equity among first and last authorship in *The BMJ* (peaking at 42% and 43%), they remain under-represented in the *Lancet* (35% and 27%). In both *The BMJ* and *Lancet*, Black authors have remained severely under-represented as both first and last authors (below 1% for most of the two decades), while Asian authors have increased proportionally to match their fraction in the general population (ranging from 2 to 10%).

Interpretation Analysis over the past two decades has shown that the gender author gap is decreasing quickly in *The BMJ* and *Lancet*. However, despite the two journals' growing focus on structural inequalities in medical academia, little progress has been made in rectifying the large gap between White British authors and other ethnic groups, especially Black authors. Without more awareness, diversity initiatives which have resulted in positive gains for White women do not seem to translate well for authors of colour.

Funding None.

Copyright © 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

eClinicalMedicine
2023;64: 102174
Published Online xxx
<https://doi.org/10.1016/j.eclinm.2023.102174>

*Corresponding author. Institute For Better Health, Trillium Health Partners - Mississauga Hospital, 100 Queensway W, Mississauga, ON L5B 1B8, Canada.

E-mail address: msa@cs.toronto.edu (M. Abdalla).

^hBoth authors contributed equally to this study.

Keywords: Ethnicity; Gender; Authorship analysis; Publication trends

Research in context

Evidence before this study

While gender equity among academic authors has been extensively investigated, there is a significant gap in our understanding of racial/ethnic authorship trends, despite the recognition of barriers to authorship along both ethnic and gender lines. Search of ((United Kingdom) AND (Authorship)) AND ((Gender) AND (Ethnicity)) in PubMed without any time restriction on January 25th 2022 returns zero relevant articles. We previously conducted an intersectional analysis of authorship in the *Journal of American Medical Association* and the *New England Journal of Medicine* that revealed unrecognized stagnation of the proportion of non-White and non-male authors.

Added value of this study

Despite growing commitments from *The BMJ* and *Lancet* to equity, diversity, and inclusion, there has been no comprehensive analysis of the impact of these initiatives on

gender and ethnic proportions among first and last authors. By analyzing two of the top UK medical journals, this bibliometric analysis reveals where initiatives for under-represented groups are making headway and where they are lacking.

Implications of all the available evidence

Over the past two decades, the proportion of women first and last authors of original research in the United Kingdom has significantly increased. However, there is a need for greater steps to be taken by the two medical journals to increase representation of non-White British first and last authors. In particular, our observations suggest that diversity initiatives have successfully improved gender representation but need to be tailored specifically to improve publications of disproportionately underrepresented ethnic groups. Diversity initiatives must be tailored specifically to benefit genders of all ethnic groups.

Introduction

The British Medical Journal (The BMJ), founded in 1840, and the *Lancet*, founded in 1823, are two of the oldest and most prominent journals in medicine. The *Lancet* has the highest impact factor of any general medical journal (impact factor [IF] of 202.7),¹ and *The BMJ* ranks fourth (IF 93.3).² Both *The BMJ* and the *Lancet* have a large and diverse readership, reaching millions of clinicians, policymakers, researchers, and patients annually.^{1,3} With extensive media coverage and impact, both journals actively influence and guide conversations among all those engaged in patient care and health policy, both in the UK and globally.

In recent years, with the goal of increasing both gender and ethnic diversity in academic authorship, the *Lancet* announced a *Group Diversity Pledge* and a *No All-Male Panel Policy* in 2018.⁴ *The BMJ* has similarly taken steps with *No All-Male Panels* and launched its first *Equity, Diversity, and Inclusion Initiatives* in 2019.⁵ Despite these endeavours to improve gender and ethnic representation, there have been few studies analysing their impact on ethnic diversity, and almost no studies that explore the intersectionality of gender and ethnicity. This is despite the fact that many studies have demonstrated the existence of diversity barriers, along both gender and ethnic lines, to academic authorship.^{6–8} Further, increasing diversity among medical students, residents, and physicians does not correspond with an increasing diversity across academic authorship in most medical specialities; in fact, the opposite trend has been noted to be true.⁹

Only one study has extensively explored gender, ethnicity, and their intersectionality in the *Journal of*

American Medical Association and the *New England Journal of Medicine*, the American counterparts to *The BMJ* and the *Lancet*.¹⁰ This study highlighted the previously unrecognised stagnation of non-male and non-White authors in recent decades in the two premiere American medical journals.¹⁰ For UK-based journals, there are no studies quantifying authorship diversity along racial and gender lines—likely due to poor meta-data collection, the lack of established methods for predicting ethnicity for UK populations, and a heavier focus on gender equity (an equally important but technically easier question to study).^{11–13}

In this manuscript, we explore both gender and ethnicity trends of all first and last authors of published articles in *The BMJ* and the *Lancet*, between 2002 and 2021. We focus on the first authorship position as it (generally) tends to reflect junior researchers/clinicians who have contributed most to the work, and the last authorship role as it is (generally) representative of established, senior principal investigators providing the most support and guidance. In addition to gender and ethnic-specific trends, we explored the limited intersection between these identities among academic authorship in the two journals.

Methods

Data extraction

Leveraging PubMed, we collected the authorship and publication metadata associated with 147,620 articles (81,529 of *The BMJ* and 66,091 of the *Lancet*). An article was defined as any document with a unique PubMed ID

(PMID), and included non-research article types, which may represent the first foray of junior or less established authors into academia. Due to changes in PubMed meta-data collection, we began our analysis from January 1st 2002, the first year in which most submissions to PubMed generally included the full forename and surname of all authors. Having both names is critical for reliably predicting race and gender for all authors.

Incomplete entries (which included, but are not limited to, incomplete/initialised names, consortia, or large group articles where no authors are explicitly listed, or lacking affiliation on all authors) were removed. We successfully selected 82,143 articles (51,209 from *The BMJ* and 30,934 from the *Lancet*), which represented >97% of all published articles published between January 1st 2002 and December 31st 2022. As we sought to understand how academic authorship reflects the diversity of the UK population, we limited our analysis to first and last authors who had a UK affiliation (18,313 articles for *The BMJ* and 5132 articles for the *Lancet*). We subsequently excluded all “news” and “comments” pieces (1577 articles for *The BMJ* and 454 articles for the *Lancet*), leaving us with 16,736 articles for *The BMJ* and 4678 articles for the *Lancet* analyses.

Statistical analysis

For our gender identity analysis, we leveraged a validated dictionary of 40,000+ first names and genders covering first names in North America, Europe (including the United Kingdom), and some overseas countries (e.g., China, India, Japan).¹⁴ The dictionary is split by country. The country includes all individuals who live in that country regardless of their place of birth or ancestry. Thus, the UK dictionary includes all African individuals living in the UK and thus the dictionary is able to provide gender identification for African, as well as all other individuals from any other continent or ancestry from any other continent, living in the UK. By running through all first names, each name gets assigned one of three categories (i.e., male [including male and mostly male names], female [including female and mostly female names], androgynous) and we later calculate the percentage of female or male authors per year.¹⁰ The numerator is the number of all male (or female) authors, including those who have published multiple times. The denominator is the total number of articles (i.e., total number of first, or last, authors, in respective analyses). As an example, suppose there are three articles published: two from one female author and one from a male author. The female percentage would be 66%. It is important to note that this dataset does not enable us to capture minority gender groups in our predictions, and therefore, following past work, we exclude names classified as androgynous and/or unknown to allow for a clearer focus on male/female author proportions.

For our ethnicity analysis, we used the Consumer Data Research Centre’s (CDRC) ethnicity estimator tool, which provides probabilistic ethnicity predictions for any given full name.^{15,16} This means a single name can be (fractionally) assigned to one or more of 12 categories: White British (WBR), White Irish (WIF), White Any Other (WAO), Asian/Asian British Indian (AIN), Asian/Asian British Pakistani (APK), Asian/Asian British Bangladeshi (ABD), Asian/Asian British Chinese (ACN), Asian/Asian British Any Other (AAO), Black/Black British African (BAF), Black/Black British Caribbean (BCA), Any Other Ethnic Groups (OXX) and unclassified. The ethnicity estimator relies on three sources of data: consumer data, the ‘Ordnance survey (Great Britain) AddressBase Premiums’ data, and the self-assigned ethnicity from the 2011 UK Census.¹⁶ All names are first geocoded by the Ethnicity Estimator to find the general location of their community/electoral area to account for differences in household census forms (this includes only unique names, i.e., each author is included only once regardless of how many times they published due to privacy limitations set by the Ethnicity Estimator). The Ethnicity Estimator then returns a probabilistic breakdown of the names that have been ethnically classified based on their forenames/surnames and indicative cultural, ethnic and linguistic origins.^{15,16} This is approximately the ethnic breakdown of unique authors; the probabilistic breakdown is an estimate to safeguard the privacy of individuals included in the database.

Importantly, the sum of percentages, both of gender and ethnic groups, do not always add up to one hundred. For gender groups, we exclude androgynous names without obviously clear gender assignments (based on census data). The ethnicity estimator does not return a fixed count for ethnic groups with fewer than 10 individuals. For years with fewer than 100 unique individuals as authors, the year was excluded from the ethnicity and intersectional gender/ethnicity analysis as the ethnicity estimator can not reliably return confident estimates. Years 2002–2008 were excluded for female first author and years 2002–2013 for female last authors for both *The BMJ* and *Lancet*.

All analyses and figures were created in R, version 4.2.1.

Ethical approval

As this analysis uses publicly available data without any participants, approval and consent was waived for this study.

Role of the funding source

This study was unfunded. There were no funders or sponsors involved in study design, data collection, data analysis or decision to publish this manuscript. S.A., Mou A., and Moh A. had full access to all the data in the study. All authors were involved in the decision to

submit the manuscript for publication. No funders were involved.

Results

Changing authorship: gender identity

Gender identity: first authors

Gender analysis of first authors in *The BMJ* reveals an abrupt decrease in estimated proportion of female authors, from 40% of all first authors in 2010, to 32% in 2011 (Fig. 1a). Female first authorship then experiences an average increase of 2.5% per year to a peak of 42% in 2021. Prior to 2010, female first authorship had averaged 34% (range 32–37%) annually. Since then, females, on average, represented an estimated 38% (range 32–42%) of all first authors and are on track to surpassing their male counterparts in the next two years if the trend continues.

In the *Lancet*, our first author gender analysis revealed a slower increase of estimated proportion of female representation among first authors. Excluding 2003, a local peak of 26%, first female authorship hovered around 18–24% from 2003 to 2007 (Fig. 1c). Since then, females represented an increasingly larger proportion of first authors, increasing on average 0.6% per

year to an all-time peak of 35% in 2017 and decreasing to 32% in 2021.

Gender identity: last authors

In *The BMJ*, the last author gender analysis paralleled that of first authors. There was a similar strong linear increase of 2.4% per year in estimated proportion of female last authors, from 29% in 2012 to 42% in 2021 (Fig. 1b). Prior to this, from 2002 to 2012, the last female authorship averaged 26% (range of 20–28%). Last female authorship in *The BMJ* reached an all-time peak of 43% in 2021, and female last authors are predicted to be on track to surpass their male counterparts in the next two years.

Our last author gender analysis in the *Lancet* paralleled that of first authors, with an almost linear increase of predicted female authorship of 0.42% per year—from 19% of all last authors in the *Lancet* in 2002 to 27% in 2021 (Fig. 1d).

Changing authorship: ethnic identity

Ethnic identity: first authors

Our ethnic identity analysis of *The BMJ* authors reveals that the estimated fraction of White first authors has remained relatively flat, from 80% in 2002 to 84% in

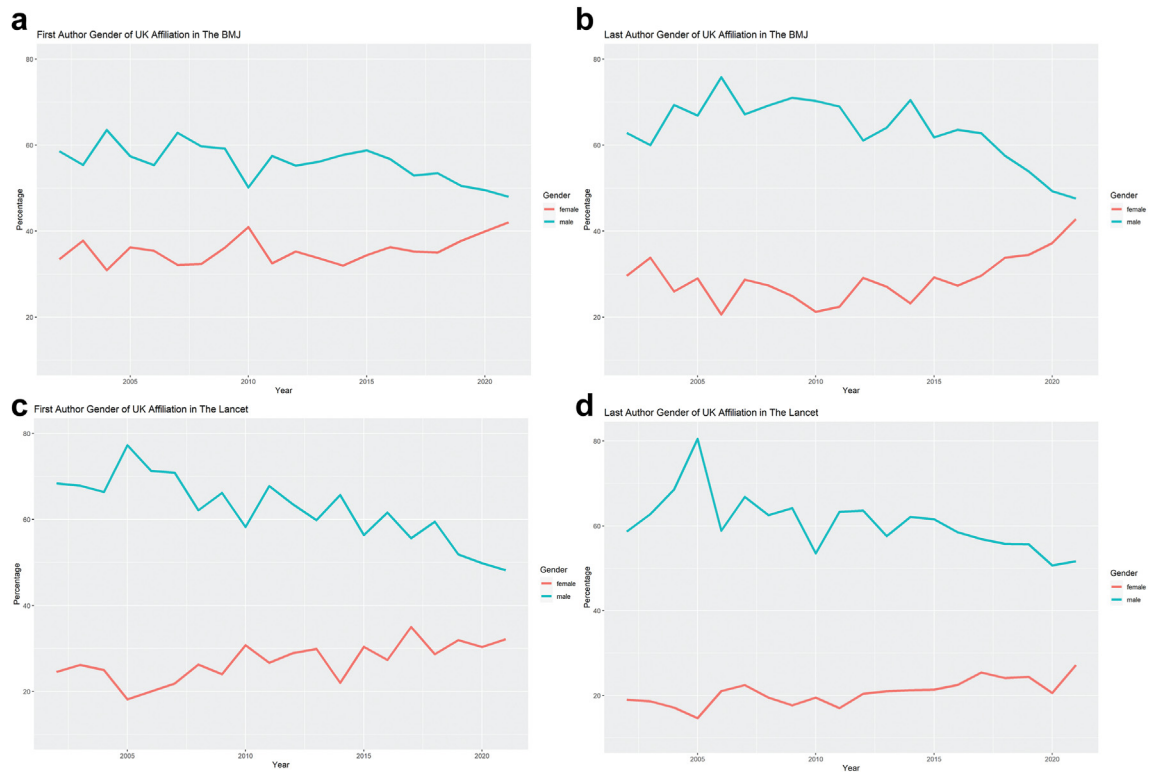


Fig. 1: Gender proportions as a function of time for first and last authors in *The BMJ* and *Lancet*. Line graph charts with gender membership as a function of year for first authors in (a) *The BMJ* and (c) *Lancet*; and for last (senior) authors in (b) *The BMJ* and (d) *Lancet*. Values do not sum to 100 as androgynous names were excluded.

2020, before dipping to 81% in 2021 (Fig. 2a). Over the past 2 decades, White individuals represented 72–88% of all first authors in *The BMJ*; at its lowest point, in 2010, this fraction never dipped below 70%. Comparatively, Asian authorship nearly doubled from 5% in 2002 to 8% in 2021. Over the past 2 decades, the estimated fraction of Asian authorship generally fluctuated between 3 and 10%. First Black authorship represented <1% from 2002 to 2003, 2005 to 2016, 2018 and 2020. Black authorship had an all-time peak of 4% in 2004 and a local peak of 2.5% in 2021. The breakdown of *The BMJ*'s White first authors (Fig. 2b) and Asian first authors (Fig. 2c) also show different levels of representation within the sub-categories.

In the above analysis, all sub-categories for White authors had sufficient data to be included (WBR, WIR, and WAO). There was sufficient data for four of the five Asian sub-categories (AIN, APK, ACN and AAO), and there was sufficient data for both Black sub-categories (BAF and BCA) (Supplementary Fig. S1).

In the *Lancet*, estimated White first authorship decreased from 84% in 2006 to 69% in 2021 (Fig. 2d). While there was substantial variance between 2002 and 2004 (range 69–80%) and 2008–2015 (range 70–81%), first White authorship consistently remained around 70.5% average after 2015. The trend for Asian authorship in the *Lancet* also differed from that observed in *The BMJ*. Despite representing 10% of first authors initially in 2002, the fraction of Asian authorship did not vary significantly over the past 2 decades (12% in 2021). Black authorship in the *Lancet* represented <2% from 2002 to 2021. The breakdown of the *Lancet*'s White first authors (Fig. 2e) and Asian first authors (Fig. 2f) also

show different levels of representation within the sub-categories.

In the above analysis, all sub-categories for White authors had sufficient data to be included (WBR, WIR, and WAO). There was sufficient data for four of the five Asian sub-categories (AIN, APK, ABD, ACN and AAO), and there was sufficient data for one Black sub-category (BAF) (Supplementary Fig. S2).

Ethnic identity: last authors

Analysis of the last author's ethnicity in *The BMJ* revealed similar results. Estimated last White authorship (WBR, WIR, and WAO) initially started with 77% in 2002, before increasing to an all-time peak of 90% in 2014, and then decreasing to 84% in 2021 (Fig. 3a). Asian last authorship has similar increases to their first author counterparts, starting at 4% in 2002 and increasing to 8% in 2021. Black last authorship in *The BMJ* was <1% from 2002 to 2013, and 2015 to 2018. It peaked in 2021 with 1.7% last authors being Black. The breakdown of *The BMJ*'s White first authors (Fig. 3b) and Asian first authors (Fig. 3c) also show different levels of representation within the sub-categories.

In the analysis for *The BMJ*'s last authors, there was sufficient data for all White author sub-categories to be included (WBR, WIR, and WAO). There was sufficient data for four of the five Asian sub-categories (AIN, APK, ACN and AAO), and both Black sub-categories to be included (BAF, and BCA) (Supplementary Fig. S1).

In the *Lancet*, the estimated percentage of White last authors had a small overall increase from 79% in 2002 to 80% in 2021 (Fig. 3d). Other than the two all-time troughs of 72% and 73% in 2006/2007, White last

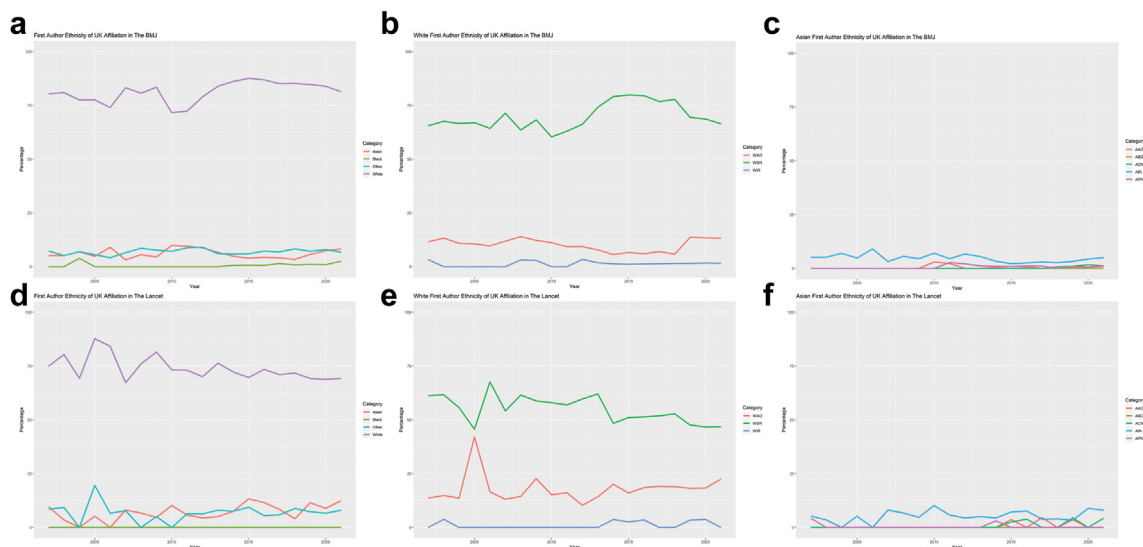


Fig. 2: Ethnicity proportions as a function of time for first authors in *The BMJ* and *Lancet*. Line graph charts of ethnicity as a function of year for first authors in (a) *The BMJ* and (d) *Lancet*; with breakdowns of White ethnicity in (b) *The BMJ* and (e) *Lancet*; and breakdowns of Asian ethnicity in (c) *The BMJ* and (f) *Lancet*. Note: In *Lancet* analysis year 2005 is excluded due to missing data hosted on MEDLINE database.

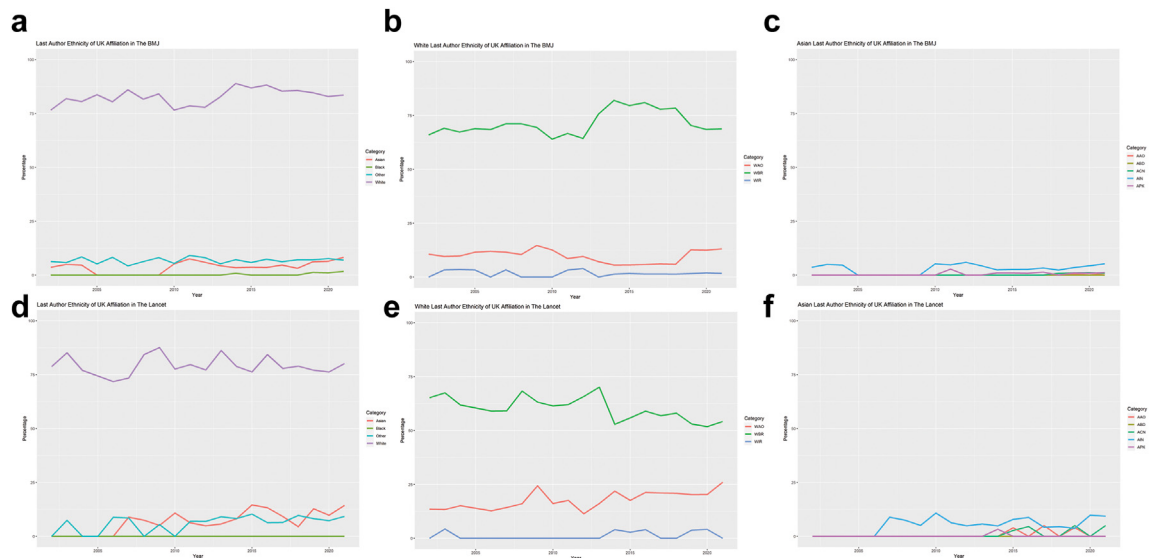


Fig. 3: Ethnicity proportions as a function of time for last authors in *The BMJ* and *Lancet*. Line graph charts of ethnicity as a function of year for last (senior) authors in (a) *The BMJ* and (d) *Lancet*; with breakdowns of White ethnicity in (b) *The BMJ* and (e) *Lancet*; and breakdowns of Asian ethnicity in (c) *The BMJ* and (f) *Lancet*. Note: In *Lancet* analysis year 2005 is excluded due to missing data hosted on MEDLINE database.

authorship steadily averaged 80% (range from 77 to 86%). Last Asian authorship generally had poor authorship rates but demonstrated significant growth. From representing <1% in 2002–2006, last Asian authorship experienced its largest single-year increase of 8.5% from 0.5% in 2005 to 9% in 2006. Asian authorship thereafter dipped below 5% only in 2012 and 2018 (4.8%, 4.4%, respectively), and slowly increased to represent 14% of last authors in 2021. Black last authors did not show the same growth and represented <1% of all last authorship from 2002 to 2021. The breakdown of the *Lancet*'s White first authors (Fig. 3e) and Asian first authors (Fig. 3f) also show different levels of representation within the sub-categories.

In the analysis for the *Lancet*'s last authors, there was sufficient data for all White author sub-categories (WBR, WIR, and WAO), all Asian sub-categories (AIN, APK, ABD, ACN, and AAO), and one Black author sub-category (BAF) to be included (Supplementary Fig. S2).

Changing authorship: intersection of ethnic/gender identity

First authors

Analysis of *The BMJ* reveals that the estimated fraction of White male first authors decreased from 45% of all first authors in 2002 to 36% in 2021 (Fig. 4a). Between 2013 and 2020, white male first authors steadily decreased at an average of –0.7% per year. This matched a nearly parallel increase of White female first authors, with an increase from 27% of all first authors in 2009 to 43% in 2015 (and subsequently fluctuating around 39% between 2019 and 2021) (Fig. 5a). We are unable to

comment on the intersection of gender and race for other ethnic groups due to insufficient numbers. There was sufficient data for two of the five Asian sub-categories (AIN and AAO) and one Black sub-category (BCA) among male first authors (i.e., >100 unique individuals). Female authors fared even worse with sufficient data for one of the five Asian sub-categories (AIN) and one Black sub-category (BAF) (Supplementary Fig. S1). Thus, we do not comment on intersectional trends of ethnic and gender groups among non-White first authors in *the BMJ*.

The intersectionality analysis in the *Lancet* revealed similar trends: White male first authors decreased from an estimated proportion of 52% in 2002 to 35% in 2021, with a proportional increase in White female first authors from 17% in 2009 to 30% in 2020 (Figs. 4c and 5c). Similar to *the BMJ*, there was sufficient data for only two of the five Asian sub-categories (AIN and APK) and neither of the Black author categories. The female authors had sufficient data for only one of the five Asian sub-categories (AIN) and one of the two Black sub-categories (BAF) (Supplementary Fig. S2). Thus, we do not comment on intersectional trends of these ethnic and gender groups among non-White first authors in the *Lancet*.

Last authors

In *The BMJ*, White male last authors had an overall decrease from 49% in 2002 to 38% in 2021 (Fig. 4b). From 2006 to 2021, White male last authorship generally decreased at –0.4% per year. White female last authors went through two sudden increases from their

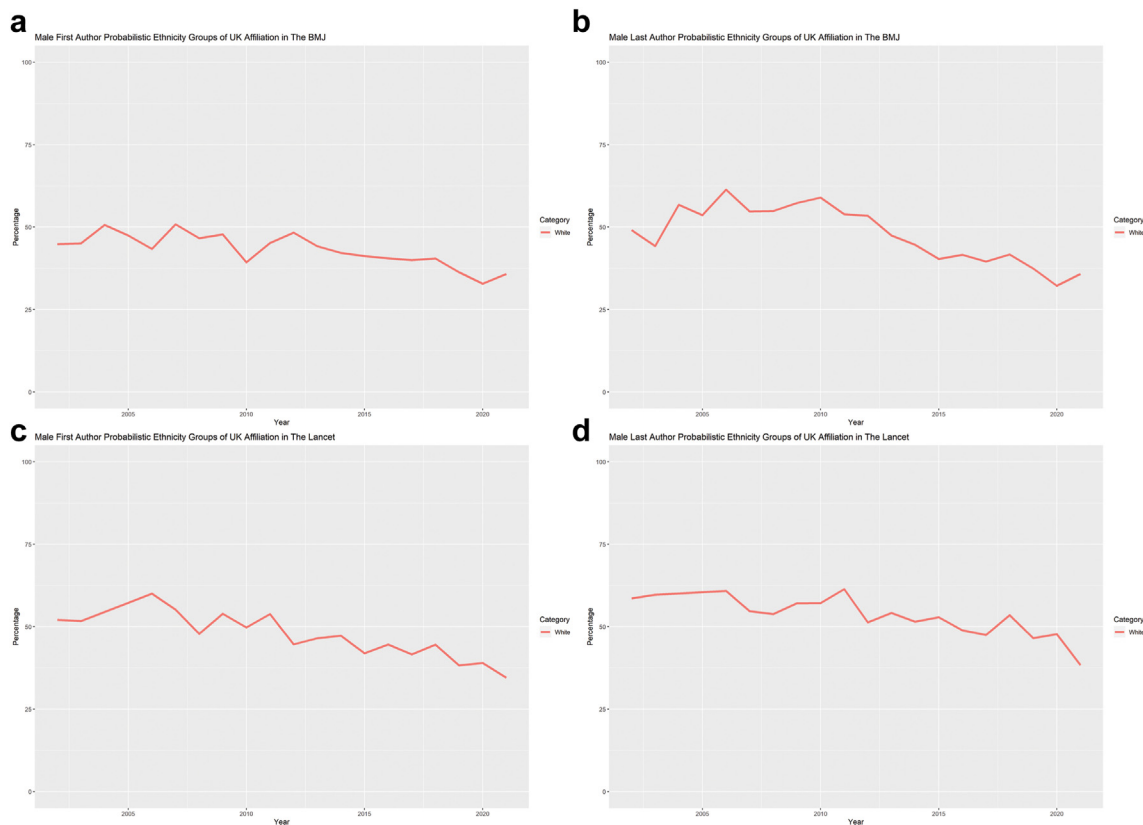


Fig. 4: Proportions of White male first and last authors as a function of time in *The BMJ* and *Lancet*. Line graph charts are a function of year for male White first authors in (a) *The BMJ* and (c) *Lancet*; and male White last authors in (b) *The BMJ* and (d) *Lancet*. Note: In *Lancet* analysis year 2005 is excluded due to missing data hosted on MEDLINE database. The two gendered ethnic values do not always sum up to their respective first and last author totals due to category exclusions and androgynous name exclusions.

troughs in 2013 (33%) and 2018 (33%) to their estimated peaks in 2015 (41%) and 2020 (40%). In 2021 White female last authors represented 40% of total authors in 2021. From 2014 to 2021, White female last authors generally fluctuated $\pm 10\%$ (Fig. 5b). We do not comment on intersectional trends of these ethnic and gender groups among non-White authors given insufficient numbers (<100 individuals in the respective categories) (Supplementary Fig. S2).

In the *Lancet*, the estimated percentage of White male last authors decreased from 58% in 2002 to 38% in 2021, although the trend was not obviously linear (Fig. 4d). The percentage of White female last authors also decreased from 24% in 2014 to 19% in 2021 (Fig. 5d). We do not comment on intersectional trends of these ethnic and gender groups among non-White authors given insufficient numbers (<100 individuals in the respective categories) (Supplementary Fig. S2).

Discussion

Our gender analysis indicates that the historical underrepresentation of woman as academic authors is being

slowly corrected. *The BMJ* is estimated to reach gender equity (defined as 50% of first and/or last authors as females) in a few years and the *Lancet* is close behind. Considering *The BMJ*'s growth only within the last five years, first female authorship is predicted to reach 50% in 2025, and last female authorship in 2024. If the *Lancet* also follows its last five years estimated growth trend, female first authorship is predicted to reach 50% in 2034, and last female authorship is expected to reach 50% in 2044.

Our ethnicity analysis, however, is more sobering. The results show that Black authors remain sorely underrepresented, and the estimated proportions have shown little improvement over the past two decades. While Black individuals represent 3.3% of the UK population, they represented less than 1% of authors for most of the two decades. Thus, improvements in ethnic and gender diversity have left out non-White (and especially Black) authors.

There are multiple limitations to our study. First, *The BMJ* and *Lancet* represent only two high-impact medical journals, and despite showing similar results to our study on *JAMA* and *NEJM*, it is not certain that such

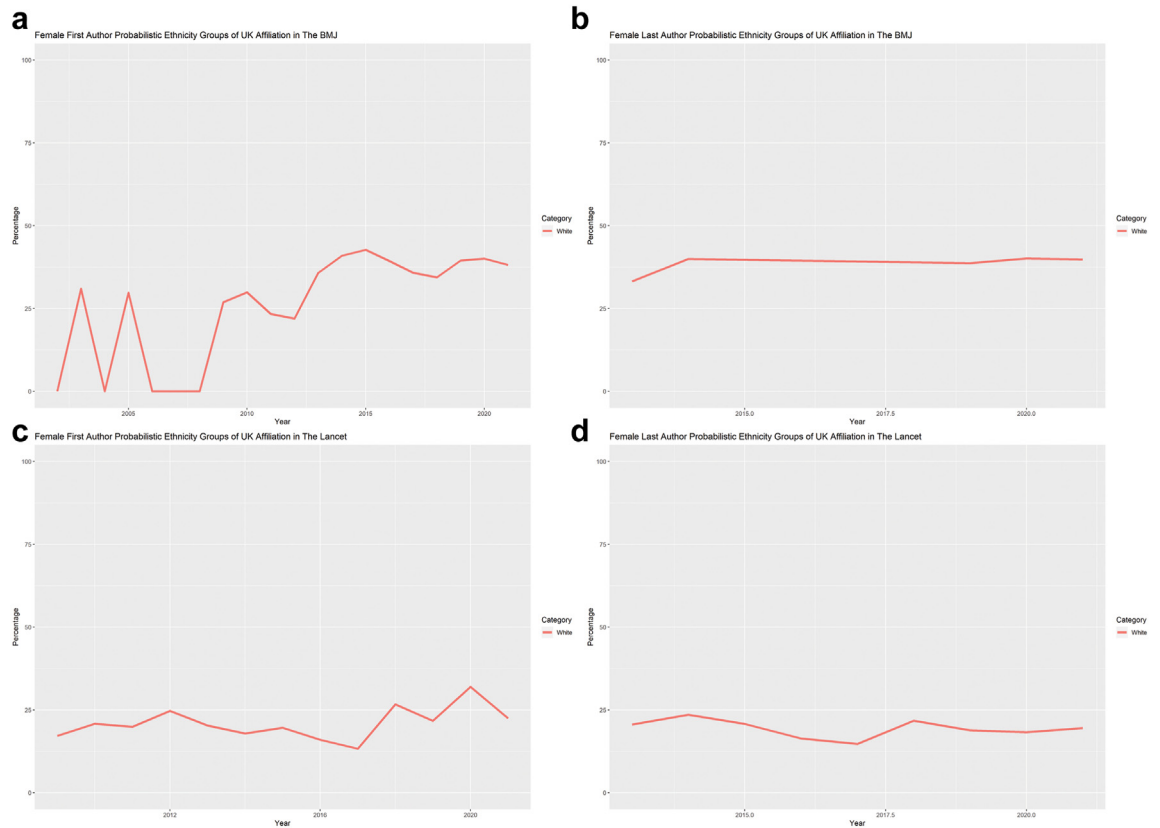


Fig. 5: Proportions of White female first and last authors as a function of time in *The BMJ* and *Lancet*. Line graph charts are a function of year for female White first authors in (a) *The BMJ* and (c) *Lancet*; and female White last authors in (b) *The BMJ* and (d) *Lancet*. Note: Female first author analysis had sufficient data from 2009+ and Female last (senior) author analysis had sufficient data from 2013+.

findings can be generalised to other journals. Second, by nature of its design, this analysis is limited to UK authors and may not reflect authorship of non-UK authors in these two journals. Third, our analysis relied on a list of publications pulled from PubMed's MEDLINE, where full forenames and surnames and affiliations were listed. All other publications (e.g., consortia papers) were excluded from our analysis. Finally, while our algorithms are peer-reviewed and validated, using names can only ever lend itself to an approximation to the proportions of race and gender without the self-identification of the authors. Our analysis fundamentally depends on a large population assumption—the estimated gender and ethnic proportions will approach the real proportions with a larger number of individuals or authors analysed. In addition, our analysis is unable to capture co-first authors and articles where the last author does not necessarily reflect corresponding author; these represent a relatively small fraction of the total number of articles.

However, the continued underrepresentation of non-White (and especially Black) authors among the two highest impact medicine journals in the UK is a reason

for concern and this analysis is the first of its kind. Our findings suggest that it is important to monitor and quantify the impact of our equity, diversity, and inclusion initiatives to ensure that all underrepresented groups benefit from these changes. Further, our observations suggest that these initiatives must be tailored specifically to benefit genders of all ethnic groups, as they otherwise often disproportionately benefit White female academics.

These authorship trends in *The BMJ* and the *Lancet* stand out in sharp contrast to those noted in our analysis of American counterparts, *JAMA* and *NEJM*, where gender equity would be centuries away given current rates of change.¹⁰ It is unclear why there is a stark difference between the two sets of journals; all four journals are single blind (i.e., the authors do not see the reviewer names/reviewers are anonymous) and have a similar peer review process. We hypothesise these differences likely reflect American versus UK differences in the academic pipeline and editorial influence as well. There is also the concern for possible "bias" against, or for, author names (or perhaps author affiliation), especially for these 4 journals. This bias against female

authors or authors from underrepresented groups may vary between journals and account, in part, for the differences we see. That said, the lack of diversity reflected from published papers might actually fare better than what exists for the paper submissions. Our gender trends also stand in contrast with those noted in a subset of medical specialities. Most notably, in academic anaesthesia research, women continue to be underrepresented as authors despite making up a substantial portion of the workforce.¹⁷ Thus, trends in *The BMJ* and *Lancet* should not be used as surrogate metrics for medical subspecialities.

Women and ethnic minorities have been historically underrepresented as authors in medical academia, and in particular, among authorship of *The BMJ* and *Lancet*.^{18–20} While there has been an influx of gender authorship analysis to explore gender trends in the UK, since 2020, the analyses tend to focus on a specific speciality (e.g., anaesthesia,¹⁷ biomedical research,²¹ or orthodontic research²²). Most have tended to select a random number of research articles from various journals in attempts to replicate general gender/geographical representation.²³ Further, none have looked at ethnicity or the intersection of gender and ethnicity, as we have done in this analysis—the first of its kind focussed on UK authors in UK medical journals.

We conclude by noting that there has been an encouraging increase in female authorship, despite predictions of plateauing progress in *The BMJ* and the *Lancet* from the 1970s.²⁴ The continual momentum of progress can perhaps be explained in part by various policies put forth by the journals, including Group Diversity pledges, No All-Male Panel Policies, and various Diversity, Equity, and Inclusion (DEI) Initiatives.⁴ However, it is important that working towards diverse academic authorship requires more than just DEI initiatives; simply increasing the proportion of racial minorities in medical schools, residencies, or in the workforce does not translate directly to increased representation elsewhere (e.g., academic authorship).⁹ We also need to formally explore additional approaches, such as double blinded review. Both the *Lancet* and *The BMJ* both are currently single blind journals. More work is needed to develop and ensure that diversity initiatives benefit genders of all racial and ethnic groups.

The continued underrepresentation of non-White (and especially Black) authors among the two highest impact medicine journals in the UK is a reason for concern. Our findings suggest that it is important to monitor and quantify the impact of our equity, diversity, and inclusion initiatives to ensure that all underrepresented groups benefit from these changes. Further, our observations suggest that these initiatives must be tailored specifically to benefit genders of all ethnic groups, as they otherwise often disproportionately benefit White female academics.

Contributors

S.A., Mou A., and Moh A. analysed the data and wrote the initial draft. All authors initiated the project, developed the idea, critically reviewed the content, revised paper, finalised manuscript, and wrote the paper. S.A., Mou A., and Moh A. had full access to all the data in the study; S.A., Mou A., and Moh A. verified the underlying data. All authors were involved in the decision to submit the manuscript for publication.

Data sharing statement

All Data collected is publicly available from NLM's annual baseline accessed as of June 2022, <https://ftp.ncbi.nlm.nih.gov/pubmed/baseline/>. Once a year, NLM produces a baseline set of PubMed citation records in XML format for download available at the aforementioned link. This provides all the data used in this analysis.

Declaration of interests

The authors declare that they have no conflicts of interest.

Acknowledgements

We acknowledge the use of the Ethnicity Estimator online tool (<https://data.cdc.ac.uk/dataset/ethnicity-estimator-software>), produced by UCL's Economic and Social Research Council (ESRC) funded Consumer Data Research Centre (CDRC), in association with the Office for National Statistics.

Appendix A. Supplementary data

Supplementary data related to this article can be found at <https://doi.org/10.1016/j.eclinm.2023.102174>.

References

- 1 *About the lancet*; 2021. <https://www.thelancet.com/lancet/about>.
- 2 History of the BMJ. <https://www.bmj.com/about-bmj/history-of-the-bmj>.
- 3 *Publishing model*; 2021. <https://www.bmj.com/about-bmj/publishing-model>.
- 4 The Editors of the Lancet Group. The Lancet Group's commitments to gender equity and diversity. *Lancet*. 2019;394(10197):452–453. <https://www.thelancet.com/diversity>.
- 5 *Equality, diversity and inclusion across BMJ company and our journals*; 2019. <https://www.bmj.com/company/edi-policy-for-bmj-journals/>.
- 6 Zeina M, Balston A, Banerjee A, Woolf K. Gender and ethnic differences in publication of BMJ letters to the editor: an observational study using machine learning. *BMJ Open*. 2020;10(12):e037269.
- 7 González-Alvarez J. Author gender in the lancet journals. *Lancet*. 2018;391(10140):2601.
- 8 Casino G, Rius R, Cobo E. National citation patterns of NEJM, the lancet, JAMA and the BMJ in the lay press: a quantitative content analysis. *BMJ Open*. 2017;7(11):e018705.
- 9 Abdalla M, Abdalla S, Maurer LR, Ortega G, Abdalla M. American Black authorship has decreased across all clinical specialties despite an increasing number of Black physicians between 1990 and 2020 in the USA. *J Racial Ethn Health Disparities*. 2023;1–9. <https://doi.org/10.1007/s40615-023-01554-0>.
- 10 Abdalla M, Abdalla M, Abdalla S, Saad M, Jones DS, Podolsky SH. The under-representation and stagnation of female, Black, and Hispanic authorship in the Journal of the American Medical Association and the New England Journal of Medicine. *J Racial Ethn Health Disparities*. 2022;10(2):920–929.
- 11 Jaggi R, Silver JK. Gender differences in research reporting. *BMJ*. 2019;367:l6692.
- 12 Rose-Clarke K, Fellmeth G. Co-first authorship and gender equity in academic publishing. *Lancet*. 2019;393(10185):2036.
- 13 Clark J, Horton R. What is the lancet doing about gender and diversity? *Lancet*. 2019;393(10171):508–510.
- 14 Michael J. *Gender-guesser 0.4.0*. 2015.
- 15 Consumer Data Research Centre. *Ethnicity estimator*; 2022. https://data.cdc.ac.uk/ethnicity_estimator.
- 16 Lan T, Kandt J, Longley P, Cheshire J, Singleton A. *Ethnicity and residential segregation*. *Consumer Data Research*. UCL Press; 2018:70–83.
- 17 Flexman AM, Shillcutt SK, Davies S, Lorello GR. Current status and solutions for gender equity in anaesthesia research. *Anaesthesia*. 2021;76(Suppl 4):32–38.

- 18 Filardo G, da Graca B, Sass DM, Pollock BD, Smith EB, Martinez MA-M. Trends and comparison of female first authorship in high impact medical journals: observational study (1994-2014). *BMJ*. 2016;352:i847.
- 19 Rexrode KM. The gender gap in first authorship of research papers. *BMJ*. 2016;352:i1130.
- 20 Lerchenmueller MJ, Sorenson O, Jena AB. Gender differences in how scientists present the importance of their research: observational study. *BMJ*. 2019;367:l6573.
- 21 Shah SA-O, Dam RA-O, Milano MJ, et al. Gender parity in scientific authorship in a National Institute for Health Research Biomedical Research Centre: a bibliometric analysis. *BMJ Open*. 2021;11(3):e037935.
- 22 Wadia R. Gender gap in authorship. *Br Dent J*. 2021;231(2):114.
- 23 Merriman R, Galizia I, Tanaka S, Sheffel A, Buse K, Hawkes S. The gender and geography of publishing: a review of sex/gender reporting and author representation in leading general medical and global health journals. *BMJ Glob Health*. 2021;6(5):e005672. <https://doi.org/10.1136/bmjgh-2021-005672>.
- 24 Sidhu R, Rajashekhar P, Lavin VL, et al. The gender imbalance in academic medicine: a study of female authorship in the United Kingdom. *J R Soc Med*. 2009;102(8):337-342.