Written Evidence Submitted by

The Royal Statistical Society, The Institute of Mathematics and Its Applications and the London Mathematical Society Committee for Women and Diversity in Mathematics

(DIV0087)

1. Introduction: diversity in mathematics, statistics and data science

This submission sets out the Royal Statistical Society's, the Institute of Mathematics and Its Applications' and the London Mathematical Society Committee for Women and Diversity in Mathematics' joint views on issues regarding equality, diversity and inclusion within the fields of mathematics, statistics and data science. Many of the issues that arise in the areas of STEM that we are concerned with are common across the mathematical sciences. So for this submission we have worked together to agree an analysis of the common problems that affect our individual disciplines and a set of recommendations intended to tackle them.

The shared analysis is set out in the introduction. In the rest of the submission, we discuss these issues from the perspective of mathematics, statistics and data science.

The Royal Statistical Society (RSS) is an academic, professional and membership organisation for statisticians and data scientists. The society represents more than 10,000 members working in academia and industry.

The London Mathematical Society (LMS) is the oldest of the United Kingdom's learned societies for mathematics. Its purpose is the advancement, dissemination and promotion of mathematical knowledge, both nationally and internationally. The society represents more than 3,000 members, nationally and internationally.

The Institute for Mathematics and its Applications (IMA) is the professional body and learned society for qualified and practising mathematicians and exists to support the advancement of mathematical knowledge and its applications and to promote and enhance mathematical culture with over 5,000 members, comprising of mathematicians from all sectors, as well as those with an interest in mathematics.

As learned societies, we promote equality, diversity and inclusion (EDI) in the field of mathematics and statistics, and the use of statistics and data analysis to promote EDI. Achieving greater diversity in the fields of mathematics and statistics is essential to the societies' goals of building a thriving cadre of mathematicians and statisticians and to using maths and statistics in the public interest. Steps to improve diversity within STEM won't just benefit underrepresented colleagues, but everyone in the field, as well as the sectors where they are employed.

1.1. Summary of key challenges in mathematics, statistics and data science

There is a lack of gender diversity among students and staff in statistics, mathematics and data science. In industry, women have higher turnover and attrition rates than men. This "leaky pipeline" has been investigated a lot and there is no single cause for it, so no silver bullet to fix it. Evidence suggests there is a feedback effect, because girls have dropped out earlier there aren't many women and because there aren't many women, women aren't attracted to the profession.

LGBTQ+ people working in STEM are more likely to experience career limitations and harassment than their non-LGBTQ+ colleagues, and women working in computer sciences and electronic engineering are more likely to have experienced bullying and harassment than male staff. This makes the field less welcoming and less attractive to a diverse set of people.

Within academia, evidence shows that women are underrepresented in senior roles. Although there has been an increase during the last decade, only 11% of UK mathematics professors are women. Evidence shows there is still a very low proportion of women studying towards a PhD in mathematical sciences compared with the proportion at undergraduate level, and there is a stagnation at mid-career, with the proportion of women declining in the transition to professorial grade.

Women submit fewer research grant applications than men and one reason for this is that they are less likely to have previously received substantial funding. The way that grant submissions are

assessed by UKRI prioritises the experience of researchers in a way that makes it difficult for less-experienced – often female – researchers to be successful in their applications. Panels of diverse people assessing applications and in the peer review process are useful and should be continued, but they alone are not sufficient – more needs to be done by UKRI to place greater emphasis on the quality of research proposals independently of the background of the researcher making the proposal.

The lack of diversity in the discipline can have an impact on society– especially in the case of algorithmic bias. This is a serious problem and it's the responsibility of those designing and developing the tools to be aware of the outcomes and correct them. Engaging a variety of stakeholders and maintaining a greater level of diversity within teams can help reduce it. It ensures a broader set of experiences are brought to bear on evaluating the ethical considerations and potential harms of these systems. Diversity of teams who are investigating algorithms should be part of the best practice.

While gender and ethnicity information are starting to be more routinely collected, socio-economic background, care leaver information, sexuality and transgender identity are not. One could argue that the lack of evidence is a result of challenges in the data that are collected on other protected characteristics (oftentimes being wildly variable or not collected at all). On top of this, different organisations have created their own definitions of STEM, and this has created a situation where estimates of diversity vary widely, and the monitoring of progress is inconsistent.

1.2. Recommendations

There are six shared recommendations that the IMA, LMS and RSS believe will help tackle these challenges:

Recommendation 1	Ensure that education in mathematics and statistics represents and reflects the interests of a broad part of society and that all students are encouraged to study mathematics and statistics by: bringing back an AS-Level qualification in maths to encourage more students to take the subject past GCSE and widening participation in Core Mathematics.
Recommendation 2	Ensure that STEM professions remain an attractive area for diverse individuals to go in to by identifying measures to tackle bullying and harassment.
Recommendation 3	Develop new educational and training routes to increase participation in data science and AI fields by students from lower socioeconomic backgrounds in consultation with relevant industries.
Recommendation 4	Develop meaningful regulatory frameworks to ensure that algorithms are not violating people's rights under the equality act, GDPR, consumer protection law or anti-competition law.
Recommendation 5	EPSRC/UKRI should introduce mid-career acceleration grants which are open to individuals who are no longer eligible for the New Investigator Award (NIA) scheme, and who have not recently applied for or held UKRI funding.
Recommendation 6	Work with professional societies to choose benchmark professions within STEM in which to track progress. For example, statisticians, data roles, engineering roles, accountancy roles, developer roles.

2. The nature and extent of underrepresentation

There is an underrepresentation of women among students studying maths-related subjects in UK Higher Education Institutions (HEI). Out of all students enrolled in mathematical sciences in the 2019/20 academic year, 37% were female students.

When looking at specific subject areas within computing and mathematical sciences the picture is varied. In the 2019/20 academic year, of the:

- 69,490 students enrolled in computer science, 19% were female students
- 3,020 students enrolled in artificial intelligence, 23% of those students were female
- 5,620 students enrolled in statistics enrollments were more balanced by gender; 45% of students were female.[1]

As well as gender diversity, there is a difference in the proportion of white students awarded a first/2:1 in mathematical sciences compared with black, Asian and minority ethnic students. In the 2019/20 academic year, 85.9% of white students were awarded a first/2:1 compared to 81.1% of black, Asian and minority ethnic students.

This difference is known as an awarding gap. In science, engineering and technology subjects – including both mathematical sciences and computer science – the awarding gap is largest between white students and black students. In the 2019/20 academic year, of the:

- 104,660 white students, 84.6% were awarded a first/2:1
- 17,080 Asian students, 78.6% received a first/2:1
- 10,010 black students, 66.2% received a first/2:1.[2]

Among academic staff across UK HEI there is also a lack of gender diversity and this has remained largely unchanged in recent years.

- In the 2019/20 academic year, there were 6,830 (34%) female staff and 13,320 (66%) male staff working in biological, mathematical and physical sciences.
- In 2014/15, these figures were 6,170 (32%) and 13,035 (68%) respectively.

Within institutions, there are more male than female staff progressing into the most senior academic roles. Research by the LMS showed that:

- In the 2017/18 academic year, 89% of mathematical science professors were male while only 11% were female. The share of women has increased from 7% in 2011/12. [3]
- The percentage of women among MSc students in mathematics has increased from 37% to 43% over the same period. However, at other stages of the pipeline the percentages of women have remained roughly constant. [4]

Women in statistics, mathematics and data science across both academia and industry have higher turnover and attrition rates than men. The Women in Data Science and AI project at The Alan Turing Institute found in the technology and IT sector, women spend almost half a year less than men, whereas the inverse is true for most other industries.[5] Some evidence points to bullying and harassment as one explanation. In a survey of postgraduate students and staff in computer sciences and electronic engineering, 49% of female staff said they had experienced harassment or bullying (compared to 34% of male staff).[6] Evidence also shows that LGBTQ+ people working in STEM are also more likely to experience career limitations, harassment, and professional de-valuation than their non-LGBTQ+ peers.[7]

3. The implications of these groups being underrepresented in STEM roles in academia and industry

Representation within the profession ensures new knowledge and insights are considered. People's backgrounds influence the evidence that they seek out and the questions they look to answer.[8] Sylvia Richardson, president of the RSS, considered this in her letter to the Times in May 2021. She highlighted the harm caused to women by the delay in the Joint Committee on Vaccination and Immunisation's (JCVI) decision to prioritise pregnant women for vaccinations against Covid-19, which wasn't helped by gender imbalances in the groups making these decisions. In December 2021 the JCVI made pregnant women a priority group for vaccination, 7 months after initial calls were made.[9]

---

[1] HESA open data and official statistics.

[2] UK domiciled first degree undergraduate qualifiers by SET category, degree class and ethnic group. Advance HE (2021) Ethnicity awarding gaps in UK higher education in 2019/20.

[3] LMS analysis of HESA data. Professors data relate to staff in mathematics cost centres in 2016/17.

[4] London Mathematics Society (2018) National Benchmarking Study 2017.

[5] Young, E., Wajcman, J. and Sprejer, L. (2021) Where are the Women? Mapping the Gender Job Gap in AI. Policy Briefing: Full Report. The Alan Turing Institute.

[6] UK Research and Innovation (2019) Bullying and Harassment in Research and Innovation Environments: An evidence review.

[7] E. A. Cech, T. J. Waidzunas (2021) Systemic inequalities for LGBTQ professionals in STEM. Sci. Adv.7, eabe0933.

[8] The Times (25 May 2021) Pregnant women need to be at front of vaccine que.

[9] UK Health Security Agency (16 December 2021) Pregnant women urged to come forward for COVID-19

One further implication of these groups being underrepresented in statistics, mathematics and data science is that it increases algorithmic bias.

As highlighted by the Alan Turing Institute:

> Over three-quarters of professionals in these [AI and data science] fields globally are male (78%); less than a quarter are women (22%) (World Economic Forum, 2018). In the UK, this drops to 20% women. This stark male dominance results in a feedback loop shaping gender bias in AI and machine learning systems.[10]

Algorithms don't perform statistically well on groups that are underrepresented in data sets used to train algorithms. For example, algorithms built using training data with an underrepresentation of black people can create outcomes that disadvantage that group. This representation bias can lead to racial discrimination if people designing and developing the algorithms are not aware of them. Engaging a variety of stakeholders and maintaining a greater level of diversity within teams can help reduce it.

In academia, one consequence of the lack of progression among female staff is a heightened difficulty of securing research funding by these individuals, which in turn functions as a barrier to progression. This is reinforced by evidence that grant submission rates by women and ethnic minority groups are lower than for white men. One explanation for why these individuals don't submit more applications is that they are less likely to have previously received substantial funding, more likely to at a lower academic rank and thus see themselves (or are seen by universities) as less likely to get funding.

In particular, the lack of targeted grant funding at mid-career level (those in permanent academic positions below professor who are no longer eligible for the New Investigator Award (NIA) scheme) hits women particularly hard, as they spend longer at mid-career and have lower funding rates at this level. The inequality in funding outcomes for mathematical sciences could be addressed by a funding scheme dedicated to mid-career scientists. It could also be addressed through protected grants. Further, UKRI could make it necessary for people writing proposals to consider diversity dimensions as well as specifically commenting on gender diversity considerations. A model for this is given by the GenderNet Plus consortium[11], and UKRI could consider joining the consortium or a similar program.

It's not just about balance of diversity on various groups but about the information they make available. UKRI could consider PI blind initial screening to ensure that when reviewers are asked to comment on the quality of the proposal they are writing about this, and not the researcher's background or experience.

The underrepresentation of black statisticians, mathematicians and data scientists can discourage others from joining. The 'Black Heroes of Mathematics' events led by the IMA President 2019-2021 showed that talented young people who sought careers advice on further study in Mathematical Sciences were encouraged to consider careers in fields where black people were more 'visible', such as sport and entertainment. Stereotyped views continued into their later careers, with difficulties progressing to PhDs despite high academic attainment at undergraduate and postgraduate levels. These experiences are aligned to many of the findings of the Hamilton Commission Report which examined the representation of black people in automotive engineering.[12]

Some diversity characteristics are less obvious. Scientists, mathematicians and engineers at a recent LGBTQ+ STEM event highlighted that they perform better in the workplace when they can be honest about their sexual identity. According to participants, not all workplaces were accepting but there were positive examples of cross-sector support networks where the UK could become a world leader for an inclusive approach to scientific research. There needs to be more training for employers across academia and industry in how to create an inclusive culture in which diversity can thrive.

4. What could and should be done by the UK Government, UK Research and Innovation, other funding bodies, industry and academia to address the issues identified

---

vaccination.

[10] Young, E., Wajcman, J. and Sprejer, L. (2021). Where are the Women? Mapping the Gender Job Gap in AI. Policy Briefing: Full Report. The Alan Turing Institute.

[11] Gender Net Plus

[12] The Hamilton Commission (2021) ACCELERATING CHANGE: Improving Representation of Black People in UK Motorsport

Considerable work has been done to improve diversity within statistics, mathematics and data science, but there is still a long way to go to overcome the underrepresentation of women, certain ethnic minorities, people identifying themselves with one of the groups in the LGBTQ+ community, people with disabilities and those from disadvantaged socioeconomic backgrounds within the profession.

Recommendation 1 Ensure that education in mathematics and statistics represents and reflects the interests of a broad part of society and that all students are encouraged to study mathematics and statistics by: bringing back an AS-Level qualification in maths to encourage more students to take the subject past GCSE and widening participation in Core Mathematics.

The underrepresentation of women in mathematics, statistics and data science professions is due to smaller numbers of women entering compared to men, and proportionately more women than men leaving at each stage in the career path. In order to make an impact, the mechanisms needed to address this should look at the whole trajectory, from primary schools to university. There needs to be advice for primary and secondary school students that enable them to see the full breadth of careers that are available within STEM, and the government should ensure that its school evaluation methods do not lead schools to discourage students from undertaking STEM qualifications because a student may not achieve top results, or may achieve better results in another subject. There should be support for professional bodies to identify and train representatives from across their networks to give talks in schools, attend STEM fairs and inspire others online.

More should be done to ensure that there is a good supply of mathematics, statistics and data science teachers. Inequalities, especially socioeconomic ones, can be addressed by providing the help that pupils need in schools rather than relying on parental support.

Recommendation 2 Ensure that STEM professions remain an attractive area for diverse individuals to go in to by identify measures to tackle bullying and harassment.

Individuals, institutions, and funding organisations all have a role to play in tackling bullying and harassment in the workplace. Measures that they can adopt include leadership training in equality, diversity in inclusion; learning more about how competition and job insecurity can contribute to bullying and harassment in the workplace; publishing the number of complaints and complaint outcomes in a standardised and transparent way.[12]

Recommendation 3 Develop new routes to increasing participation in data science and AI fields by students from lower socioeconomic backgrounds, in consultation with relevant industries.

University study is not the only route into statistics and data science professions. Non-university routes such as degree apprenticeships could help encourage a broader group of people into the statistics and data science professions. This would particularly help those from lower socioeconomic backgrounds, for whom the costs associated to university study and less community and family experience with university create barriers to university study. In 2020 research by the Government's Social Mobility Commission showed that the potential benefits of apprenticeships are not being felt by student from disadvantaged backgrounds. The number of students from disadvantaged backgrounds had declined by 36% since the introduction of the Apprenticeship Levy.[13]

While degree apprenticeships may support people into wider roles in data (for example, data technicians, data engineers and data QC/QA engineers), there needs to be a higher level of statistics and maths taught to underpin knowledge in data science and AI. Government should also ensure that employers are supported in being able to develop Apprenticeship Standards more quickly, which align to professional standards.

Recommendation 4 Develop meaningful regulation and best practice to ensure ethical regulation and development of algorithms with diversity of teams working on development implementation and regulation being a main aspect of good practice.

---

[12] Seven principles to change the UK's research culture

[13] Social Mobility Commission (2020) Apprenticeships and social mobility.

Algorithmic bias can discriminate against protected characteristics and regulatory frameworks should be established to ensure that when used in decision making, they are not violating people's rights under the equality act, GDPR, consumer protection law or anti-competition law. The differential impact on people could be guarded against by observing recommendations in current literature, such as the IEEE P7003 standard for algorithmic bias considerations. Possible frameworks include: diversity of teams who are creating algorithms; benchmarking procedures and criteria for the selection of validation data sets for bias quality control; guidelines on establishing and communicating the application boundaries; suggestions for user expectation management.[13] Such a framework could be shaped in conjunction with the new Data, Technology and Analytics (DaTA) unit at the Competition and Markets Authority (CMA).

Recommendation 5 EPSRC/UKRI should introduce mid-career acceleration grants which are open to individuals who are no longer eligible for New Investigator Award (NIA) scheme and who have not recently applied for or held UKRI funding.

The proportion of the professoriate who are women has increased but gender parity is still a long way off. There is a need for more funding support for mid-career researchers (those in permanent academic positions below professor who are no longer eligible for the New Investigator Award (NIA) scheme). Such a scheme would support individuals who reached mid-career before protected early career support was introduced. It would support talented scientists who have had a period of demanding caring work or work-related responsibilities -- for example those taken on during the Covid-19 pandemic – and enable them to refocus on research. This would benefit mid-career women, but also ethnic minorities, who have experienced similar disadvantages, and men with caring responsibilities who have taken on large administrative burdens.

Recommendation 6 Choose benchmark professions within STEM in which to track progress e.g. data science, statistics, data roles, engineering roles, accountancy roles, developer roles.

Different organisations have created their own definitions of STEM, and this has created a situation where there are widely varying estimates of diversity within STEM. Issues are also apparent with the definition and categorisation of EDI data. Many employers don't collect demographic information consistently, and the information being collected is not always relevant for addressing emerging diversity concerns. There is a need for much more detail in the available statistical information concerning protected characteristics and underrepresented groups.

There is often a lack of meaningful data to discern whether some groups are underrepresented – such as LGBTQ+ groups, individuals from socioeconomic disadvantaged backgrounds and care leavers. Unless some measure of agreement is reached on the statistical information that underpins the concerns about diversity, progress is difficult to measure.

(January 2022)

[13] A. Koene, L. Dowthwaite, and S. Seth (2018) IEEE P7003 Standard for Algorithmic Bias Considerations. FairWare'18, May 29, 2018, Gothenburg, Sweden. 4 pages. https://doi.org/10.1145/3194770.3194773