




BMJ Open Cross-sectional Survey of Medical student Attitudes to Research and Training pathways (SMART) in the UK: study protocol

Sophie Roche , Soham Bandyopadhyay, Alexander Grassam-Rowe, Robin Andrew Brown , Poppy Iveson, Garry Mallett , Holly Eggington, Catherine Swales

To cite: Roche S, Bandyopadhyay S, Grassam-Rowe A, *et al.* Cross-sectional Survey of Medical student Attitudes to Research and Training pathways (SMART) in the UK: study protocol. *BMJ Open* 2021;**11**:e050104. doi:10.1136/bmjopen-2021-050104

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2021-050104>).

Received 11 February 2021
Accepted 19 July 2021



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY. Published by BMJ.

Medical Sciences Division, University of Oxford, Oxford, UK

Correspondence to

Dr Sophie Roche;
sophie.l.roche@googlegmail.com

ABSTRACT

Background An understanding and appreciation of scientific research is a key quality of the modern clinician. Yet the Medical Schools Council has previously reported a reduction in the number of clinicians performing research. To explore the reasons for this difficulty, this multicentre, cross-sectional study aims to determine the medical student involvement and perceptions of research and research-orientated careers. It will additionally identify perceived barriers and incentives to participating in research as a student.

Methods and analysis This cross-sectional study of medical students at UK medical schools recognised by the General Medical Council will be administered using an online questionnaire. This will be disseminated nationally over a 2-month period through collaborative university medical school and student networks. The primary outcome is to determine the extent to which medical students are currently involved in research. Secondary outcomes include identifying the personal and demographic factors involved in incentivising and deterring medical students from becoming involved in research during medical school. This will be achieved using a selection of Likert scale, multiple-choice and free text questions. Ordinal logistic regression analysis will be performed to understand the association between specific factors and student involvement in research. This study will also characterise the proportion of medical students who are currently interested in conducting research in the future.

Ethics and dissemination Ethics approval has been obtained from the Medical Sciences Interdivisional Research Ethics Committee, Oxford, England. The results will be disseminated via publication in a peer-reviewed medical journal and may be presented at local, regional, national and international conferences by medical student collaborators.

INTRODUCTION

It is of vital importance to the progress of medical innovation for clinicians to perform research as their combined academic and clinical perspectives allow the most pressing challenges in healthcare to be identified and

Strengths and limitations of this study

- We will be the first to examine the attitudes of current UK medical students towards previous and future research opportunities.
- The design is limited by its snapshot nature, as we are unable to see how attitudes have or will change over time.
- It is possible that those already interested in this area will be more likely to complete the survey.

studied. Their position affords an integrative outlook, evaluating practice objectively while advancing our collective knowledge.¹ Clinicians playing an active role in research often ensure that advances made in that research are incorporated into clinical practice as efficiently as possible.²⁻⁴

While it is difficult to quantify the total number of clinicians involved in research, the total number of clinical academic staff employed by UK medical schools has declined since its peak in 2010.⁵ An insufficient number of individuals entering the clinical academic 'pipeline' have been identified as a contributing factor in this downward trend,⁵ resulting in reduced replacement of an ageing workforce. This is surprising given that involvement in research is often seen as a desirable quality in the Curriculum Vitae (CV) of medical students and junior doctors.⁶ Indeed, one survey found career success was independently associated with having conducted research as a student.⁷ Furthermore, it has been shown that intercalation can have positive academic consequences.⁸ Additionally, it is a General Medical Council (GMC) requirement for students to have a working knowledge of research.⁹ It is therefore important to gauge the aspirations and career goals of students, as well as their



attitudes towards careers which involve research and clinical academia.

While there have been attempts to survey the barriers and enablers of progression for early career clinical academics,^{10 11} there are currently no published systematic surveys of medical students which seek their views on doing research both as a student and during their career as a doctor. Furthermore, recent changes to the centralised application system have meant that research publications and additional degrees are no longer granted points as of 2023,^{12 13} despite such points having been shown to be one of the primary motivating factors for medical student involvement in research.^{14 15} This change has proved controversial^{13 15} and several authors have raised concerns that fewer medical students will now become involved with research.

Given the need to maintain, and indeed increase, the number of clinicians involved in research, it is important to identify the factors influencing medical students in deciding whether to engage in research. Previous studies have identified intrinsic demographic factors which may affect interest in particular specialty training programmes, including academic training pathways, and research, such as economic background, ethnicity and gender.^{16 17} Non-demographic influences have also been identified, including previous involvement in research and access to mentors.^{16–18} Similar data are however currently lacking for medical schools in the UK, and no clear data available for the interactions between demographic and non-demographic influences. With variation in both medical school requirements for involvement in research during training and opportunities to intercalate,¹⁹ it is important that we understand the different attitudes towards clinical academia among UK medical students. We also hope to describe the experiences that have conceivably shaped these, and the perceived barriers to pursuing research during medical school and beyond.

The Survey of Medical student Attitudes to Research and Training pathways (SMART) study is an online national questionnaire-based study. The aim of this study is to ascertain medical student perceptions of research and research-orientated careers, including perceived barriers and incentives to participating in research as a student and in following a research-orientated career. It is our hope that increasing awareness of the issues medical students face will encourage solutions to be sought by medical schools and regulators. SMART will also identify key issues for the Academy of Medical Sciences to address as part of their long-term INSPIRE strategy.²⁰

METHODS AND ANALYSIS

Primary aim

- ▶ To determine current medical students' involvement with research.

Secondary aims

- ▶ To characterise the group of students who are interested in research.
- ▶ To identify the factors that drive medical students to conduct research.
- ▶ To identify the perceived barriers that have historically prevented and continue to prevent current medical students from partaking in research.
- ▶ To determine if current medical students are interested in conducting research.
- ▶ To identify factors that could encourage current medical students to conduct research.

Study design

SMART is an online, national, multicentre, questionnaire-based study focusing on medical student perceptions of research and research-orientated careers, including barriers and incentives to participating in research as a student and in following a research-orientated career. The questionnaire will be disseminated through collaborative university medical school and student networks, such as the network of INSPIRE leads across the country. It is well documented that the generic collaborative method works, and that participating students benefit from involvement.^{21–26} In short, the collaborative method involves a 'snap-shot, protocol-driven, pragmatic multi-centre research'²⁴ (p 355), approach undertaken by separate groups of trainees, or in this case students; the collaborative method also allows for greater size and power of studies.²¹ The SMART study will be delivered by a team of University of Oxford medical students in clinical years 4, 5 and 6. The questions included can be found in online supplemental appendix S1. The questions were formulated by medical students at Oxford University. A brief review of the existing literature was performed to identify the gaps in knowledge and to also look at similar questionnaires and qualitative studies on the viewpoints of students and academics. This allowed an understanding of domains and items relevant to determining the aim of the project. Two students separately suggested questions which were pooled before being considered by all authors. Medical student and academic staff feedback was also sought at this point. The questionnaire has face and content validity, there is no gold standard to compare against for criterion validity. Questions were adjusted following reviewer comments to make them as non-directive as possible. Construct validity will be checked during the data collection period. Feedback will be collected from a pilot study regarding the suitability of questions, length of questionnaire and suggestions for other aspects to consider.

Patient and public involvement

No patients or members of the public were involved in the design of the study. We will share the results with interested individuals and publicly via journal publication.

Table 1 Project timeline

Dates	Activity
1 April 2021 to 30 May 2021	Recruit collaborators at all UK universities via collaboration with INSPIRE student leads
1 September 2021 to 1 October 2021	Study pilot to be run at Oxford
1 October 2021 to 30 November 2021	Study to be modified based on pilot feedback.
1 September 2021 to 1 October 2021	Study set-up, for example, training collaborators, gaining approval for this study at each centre, providing collaborators with Qualtrics logins.
1 November 2021 to 1 January 2022	Study runs nationally for 2 months.
1 January 2022 to 31 March 2022	Analyse data and prepare manuscript.

Extended data collection periods may be incorporated to grant flexibility to centres that may have experienced logistical obstacles to study commencement.

Methods for recruiting participants

All medical students currently studying for a UK medical degree at a UK medical school recognised by the GMC will be eligible to participate. A list of these medical schools can be found in online supplemental appendix S2.

Medical students will be invited to participate in the study through several routes:

- ▶ Medical societies.
- ▶ Medical school mailing lists.
- ▶ Organisations focused on academic medicine, such as the Academy of Medical Sciences.
- ▶ Conferences.
- ▶ Freshers' Fairs.
- ▶ Social media.

In addition, medical students enrolled in the medical schools in online supplemental appendix S2 will be invited to collaborate in the study as regional leads as described in table 1. The maximum number of collaborators from each medical school will be 1 per year. Online supplemental appendix S3 details the participant facing information used in recruitment for the SMART study.

Collaborators will ensure that their medical school is formally engaged at an early stage of this study, and they will be primarily responsible for disseminating this questionnaire among students at their medical school. Medical school collaborators will be able to request for their own specific data and the analysis done on said data from the SMART steering committee following study completion. These data will be anonymised. Researchers will have access to these anonymous data.

Information provided to participants

The following pieces of information will be provided to participants before taking part in the study. It will be

attached to recruitment emails, and appear as the front page of the questionnaire:

- ▶ Name of the study: Cross-sectional Survey of Medical student Attitudes to Research and Training pathways (SMART) in the UK .
- ▶ Name of the principal researcher carrying out the study and information on how to contact him: Soham Bandyopadhyay, soham.bandyopadhyay@st-hildas.ox.ac.uk.
- ▶ *What is the aim of this study?* This study aims to ascertain current medical student involvement with research. We also hope to identify factors encouraging and discouraging students from partaking in research and to consider what may encourage more engagement with scientific research in the future.
- ▶ *Why have I been selected to take part?* You are being invited to take part in the questionnaire as you are a medical student currently studying for a UK medical degree at a UK medical school recognised by the GMC.
- ▶ *What do I have to do?* If you choose to participate in this voluntary survey, you will be asked to complete a questionnaire about your background, your previous exposure to research and your feelings towards a research career. This study is voluntary. If you decide not to participate this will not impact your academic standing in any way. If you decide to take part, you will be asked to complete the survey by clicking on the link below. This survey is expected to take about 10–15 min to complete, but there is no time limit and you can take as much time as you like. No background knowledge is required. We will ask for your consent for the collection and storage of data in accordance with the UK General Data Protection Regulation (GDPR) within the survey. For more information on GDPR please click on the following link: <https://gdpr-info.eu/>.
- ▶ *Do I have to participate?* Please note that your participation is voluntary. You may withdraw at any point during the questionnaire for any reason, before submitting your answers, by closing the browser. In cases of withdrawal from the study, no new data will be collected or linked to other data from that point on. If you do not want to answer some of the questions you do not have to, but you can still be in the study. All questions are optional. Your decision whether or not to be part of the study will not affect your academic standing or your access to university support services. If you have already submitted data and wish to withdraw from the study, please contact soham.bandyopadhyay@st-hildas.ox.ac.uk by 1 October 2021.
- ▶ *Who has approved this study?* This project has received ethics clearance through the University of Oxford's ethical approval process for research involving human participants (reference R73479/RE001).
- ▶ *How will my data be used?* Your answers will be completely anonymous, and we will take all reasonable measures to ensure that they remain confidential. Your data will be stored in a password-protected



file and may be used in academic publications. Your IP address will not be stored. If you provide us with your email address, we will delete that information at the end of the study. No answers will be linked to your email address. Research data—your anonymised answers—will be stored for a minimum of 10 years after publication or public release.

- ▶ *Who will have access to my data?* Qualtrics is the data controller with respect to the personal data they hold about you and, as such, will determine how your personal data are used. Please see their privacy notice here: <https://www.qualtrics.com/privacy-statement>. Qualtrics will share any email address you provide and your anonymised answers with the University of Oxford, for the purposes of research. Researchers involved in the project will have access to these anonymised data. The University of Oxford is the data controller of university email addresses, please see their privacy notice here: <https://complianceadmin-ox.ac.uk/student-privacy-policy>. Responsible members of the University of Oxford and funders may be given access to data for monitoring and/or audit of the study to ensure we are complying with the guidelines, or as otherwise required by law.
- ▶ *Are there any benefits to taking part?* Despite not have any immediate individual benefits by participating in this survey, you are given the opportunity to contribute to valuable and innovative research which could be used in the future by medical universities and the world. You may find this survey an opportunity to self-reflect. There will be the option to submit email address in order to be entered into a prize draw. At the conclusion of data collection, two random participants will be awarded £50 in Amazon vouchers, two further random participants will be awarded £25 in Amazon vouchers. This will be optional, as it requires you to provide personally identifying data (ie, contact details). These will not be linked to the questionnaire answers given, and will only be used for contact regarding relevant rewards as above.
- ▶ *Will the research be published?* The findings of the study may be published in peer-reviewed journals, presented at relevant conferences and meetings and a summary of the findings will be made available on social media.
- ▶ *Are there any possible risks involved with my participation?* Some of the questions that we ask may cause upset. If you experience any distress from participating in this study, you may stop the survey at any time or skip any upsetting questions. If your distress continues after leaving the survey, we have provided a list of supportive services nationwide that can be helpful and that you might consider contacting (online supplemental appendix S4, to be linked here, and appear again at the close of the survey).
- ▶ *Who do I contact if I have a concern about the study or I wish to complain?* If you have a concern about any aspect of this project, please speak to the researcher (SB) at soham.bandyopadhyay@st-hildas.ox.ac.uk who will do

his best to answer your query. The researchers should acknowledge your concern within 10 working days and give you an indication of how they intend to deal with it. If you remain unhappy or wish to make a formal complaint, please contact the Chair of the Medical Sciences Interdivisional Research Ethics Committee. Email: ethics@medsci.ox.ac.uk. Address: Research Services, University of Oxford, Wellington Square, Oxford OX1 2JD OR. The Chair will seek to resolve the matter in a reasonably expeditious manner.

- ▶ *How do I find out what was learnt in this study?* This study is expected to be completed by approximately March 2022. If you would like a brief summary of the results, please write to us by email to request information.
- ▶ *Who to contact for further details?* For any further questions or more information on the study, please contact us on the following email address: soham.bandyopadhyay@st-hildas.ox.ac.uk. Alternatively, you could contact the principal investigator (CS) at catherine.swales@ndorms.ox.ac.uk.

A representation of how this information will be presented to medical students can be found in online supplemental appendix S5.

Financial and other rewards to participants

A prize draw involving participants who have opted to provide their contact details will take place at the conclusion of data collection. Two random participants will be awarded £50 in Amazon vouchers. Two further random participants will be awarded £25 in Amazon vouchers.

Data collection

The online questionnaire consists of 23 quantitative and qualitative questions that use a combination of the Likert scale, multiple-choice options and free text in order to broaden the capture of sentiment nuance and improve precision in the data. The questionnaire has been sent to medical students who were not involved in creating the data collection proforma. Their responses were evaluated for potential problems, and the questionnaire was updated to ensure the questions were relevant, comprehensive and accessible. A pilot study will be performed at Oxford Medical School before this study is launched nationally. There are 952 students enrolled at Oxford Medical School. Data collection will take place at the beginning of September 2021, aiming to be complete by November 2021. Based on the experiences from the pilot, the questionnaire will be modified to improve clarity, objectivity and acceptability (online supplemental appendix S6).

Primary outcome

To describe the extent of current medical students' involvement with research.

Secondary outcomes

To identify and understand the reasons driving and excluding medical students from research. To identify

factors that might encourage more current students to conduct research.

Data management

All the data will be anonymised and stored in Qualtrics. Qualtrics is a cloud-based platform, with the ability to create and customise databases. Qualtrics guarantees the highest levels of security for stored data and is compliant with regulations including GDPR and ICH E6 Good Clinical Practice.²⁷ The research team will have individual accounts, with access to data determined by the study administrator. Audit logs are also available to provide an overview of data access and modifications.

Statistical analysis

We will use descriptive statistics to achieve our primary outcome in describing the extent of current medical students' involvement with research.

In order to achieve our secondary outcomes we will consider the following research questions using machine learning techniques to predict in our cohort:

Q1: What factors are associated with whether an individual in our cohort intends to pursue an academic career?

Q2: What factors are associated with a respondent's interest in undertaking (more) research in the future?

Statistical procedure

Unless explicitly stated otherwise, assume $\alpha=0.05$, and $CI=95\%$. Data will be cleaned of any conflicting, impossible or corrupted data by deletion of all entries of the same individual.

Data processing for descriptive statistics

This section seeks to provide an easy to interact with description of the responses to our survey, and the demographics of those who responded.

1. Percentages for responses by each question.
2. Median and deciles of spread for answers to appropriate question(s).
3. Mode for answers to appropriate question(s).
4. Number of missing values per question. Data will be assumed to be missing completely at random.
5. Number of individuals with one or more missing values.
6. Identify sampling weights for demographic populations using data from local sources, and national data collected about medical student demographics at each school.
7. Key qualitative themes identified with free writing texts of questions 15, 16, 17, 23. We will look to identify common sentiments, having trained multiple different team members to independently look across the data and identify recurrent themes and sentiments. Each response will be evaluated by at least two of the trial team.

Establishing correlation across responses

The tests described in this section aim to find any predictive associations and identify likely explanatory variables.

As a method, it also allows capture of the entirety of the cohort data.

We will perform multiple correspondence analysis using sampling weights from 1.5. We will account for ordinality in data with orthogonal polynomials,²⁸ where appropriate, to produce a simple visualisation of correlation across variables. All interval/ratio variables will be appropriately binned to produce ordinal variables for the analysis above. If there are a low number of responses then answers to question 4 may be binned into broader ethnicities and backgrounds in order to maintain statistical power.

Testing Q1: what factors are associated with whether an individual in our cohort intends to pursue an academic career?

We will run an ordinal logistic regression using the cumulative logit link function. We will use the ordinal responses to question 20. 'How much do you agree with the statement: "I wish to pursue an academic career or an academic training pathway".' (Likert scale of 'Disagree' through to 'Strongly agree') as our response variable, and use as predictor variables the answers to other questions. Detail on proposed model construction can be found in online supplemental appendix S7.

Testing Q2: what factors are associated with a respondent's interest in undertaking (more) research in the future?

The methods will be the same as above using the ordinal responses to question 22. 'How strongly do you agree with the statement: "I would be interested in undertaking (more) research in the future".' (Likert scale of 'Disagree' through to 'Strongly agree') as our response variable, with answers to all questions as predictor variables as detailed in online supplemental appendix S8 model construction.

Subgroup analysis

With a priori assumptions about significant contributors to variance, one may attempt to control for these through subgroup analysis, in categorisation of the data points by a given variable or variables.

However, to ensure validity of causal inference, one must ensure that the appropriate covariates are selected for any regression model attempting to interrogate such data.²⁹ We believe a priori that we may see a large degree of variance between those who have already completed research and those who have not. To account for this, we propose two models (online supplemental appendix S9, equations 3 and 4) to account for this difference and ascertain key contributors in these different groups, with the goal of more policy-relevant inferences being drawn for each subgroup.

Additional subgroup analysis may be carried out if key contributors to variance are identified in the multiple correspondence analysis, or otherwise. Additional exploratory analysis may be carried out as deemed necessary, but it will be reported as such, and will be carried out with good statistical reporting in mind.



Authorship

In accordance with the National Research Collaborative authorship guidelines,²⁴ all publication outputs from SMART will be listed under a unified corporate authorship: 'SMART Collaborative'. Certain publications will include named authors on the byline as well as the group name. This will follow the example set by commendable collaboratives including STARSurg and InciSioN UK.^{25 26} Anyone who has demonstrated satisfactory completion of the minimum requirements for authorship will be eligible for PubMed-citable collaborative authorship in accordance with the roles defined below.

Writing group

Responsible for the overall scientific content, data analysis and preparation of research manuscripts.

Steering committee

Responsible for the protocol design, project coordination and data handling.

Collaborators

A network of medical students across all medical schools. They are responsible for leading the study regionally.

ETHICS AND DISSEMINATION

Ethics

Ethics approval has been obtained from Medical Sciences Interdivisional Research Ethics Committee, Oxford, England (reference R73479/RE001).

Dissemination

The protocol will be disseminated primarily through recruited medical student collaborators. Should UK medical schools wish to see the protocol, collaborators may pass it along as well. Any publications of the protocol will be advertised through social media.

Following study completion, teleconferences will be held with all collaborators to share and discuss the data analysis undertaken and the study results. Following this, the results will be presented at local, regional, national and international conferences by medical student collaborators. A standard PowerPoint presentation and poster will be created for this purpose. All presentations will be coordinated by the SMART steering committee to avoid duplications and to ensure all conference regulations are fulfilled. In addition, the results will be disseminated via publication in a peer-reviewed medical journal. All collaborators will be given PubMed-citable collaborative coauthorship under the institutional name 'SMART Collaborative'. We will have a hybrid authorship list of named authors and the institutional collaborative.

Following publication, the manuscript can be shared by collaborators with their medical schools to feedback the study results, and to highlight the scope for expanded integration of research within medical school curricula. Medical schools can request for their own specific data and the analysis done on said data from the steering

committee following study completion. The fully anonymised data set will be made publicly available.

Twitter Holly Eggington @HollyEggington

Contributors The study concept and design was conceived by SB and SR. Abstract was written by GM, PI and SR. Strengths and limitations were written by SR. Introduction was written by SR, RAB and HE. Methods were written by SR. Statistical analysis was written by SR and AG-R. AG-R provided the statistical planning and analysis, and contributed to survey design. Authorship was written by SR. Ethics and dissemination was written by SR. CS and SB had a supervisory role at all stages of manuscript preparation, and helped prepare and write funding applications. All authors critically reviewed the manuscript.

Funding This work was supported by an INSPIRE grant from the Wellcome Trust.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution 4.0 Unported (CC BY 4.0) license, which permits others to copy, redistribute, remix, transform and build upon this work for any purpose, provided the original work is properly cited, a link to the licence is given, and indication of whether changes were made. See: <https://creativecommons.org/licenses/by/4.0/>.

ORCID iDs

Sophie Roche <http://orcid.org/0000-0001-6751-0636>

Robin Andrew Brown <http://orcid.org/0000-0003-2999-884X>

Garry Mallett <http://orcid.org/0000-0001-5475-9061>

REFERENCES

- Jacob H. Why all doctors should be involved in research. *BMJ* 2016;352:i164.
- Rees M, Dangerfield P, Katz D. *Every doctor a scientist and a Scholar*. BMA Publications, 2015.
- BMA. *The role of the clinical academic*. BMA Publications, 2014.
- Savill J. More in expectation than in hope: a new attitude to training in clinical academic medicine. *BMJ* 2000;320:630-3.
- Medical Schools Council. Survey of medical clinical academic staffing levels, 2017. Available: www.medschools.ac.uk/SiteCollectionDocuments/Survey-Medical-Clinical-Academic-Staffing-Levels-2017.pdf
- Chambler AF, Chapman-Sheath PJ, Pearse MF. A model curriculum vitae: what are the trainers looking for? *Hosp Med* 1998;59:324-6.
- Brancati FL, Mead LA, Levine DM, et al. Early predictors of career achievement in academic medicine. *JAMA* 1992;267:1372-6.
- Mahesan N, Crichton S, Sewell H, et al. The effect of an intercalated BSc on subsequent academic performance. *BMC Med Educ* 2011;11:76.
- General Medical Council. Outcomes for Graduates [online]. Available: <https://www.gmc-uk.org/education/standards-guidance-and-curricula/standards-and-outcomes/outcomes-for-graduates/outcomes-for-graduates> [Accessed May 2021].
- Lopes J, Ranieri V, Lambert T, et al. The clinical academic workforce of the future: a cross-sectional study of factors influencing career decision-making among clinical PhD students at two research-intensive UK universities. *BMJ Open* 2017;7:e016823.
- Oldfield K, Turner C, Garnett E. A Cross-Funder review of early-career clinical academics: Enablers and barriers to progression. Available: <https://mrc.ukri.org/documents/pdf/review-of-early-career-clinical-academics/> [Accessed 12 May 2021].
- Medical Schools Council. MSC statement on the UKFPO decision to remove educational achievements from the foundation ranking

- process. Available: <https://www.medschools.ac.uk/news/msc-statement-on-the-ukfpo-decision-to-remove-educational-achievements-from-the-foundation-ranking-process> [Accessed 12 May 2021].
- 13 British Medical Association. Additional achievements ruled out of foundation programme applications [online]. Available: <https://www.bma.org.uk/news-and-opinion/additional-achievements-ruled-out-of-foundation-programme-applications> [Accessed 12 May 2021].
 - 14 Pang KH, Hobbis C, Burleigh EJ, *et al.* Publication component of the UK Foundation programme application: perception of medical students. *Adv Med Educ Pract* 2020;11:735–40.
 - 15 McVeigh J, Jeilani M, Super J. Factors influencing the engagement in research of medical undergraduates and clinical trainees. *Postgrad Med J* 2021;postgradmedj-2021-139793.
 - 16 Mina S, Mostafa S, Albarqawi HT, *et al.* Perceived influential factors toward participation in undergraduate research activities among medical students at Alfaisal University–College of medicine: a Saudi Arabian perspective. *Med Teach* 2016;38(Suppl 1):S31–6.
 - 17 Amgad M, Man Kin Tsui M, Liptrott SJ, *et al.* Medical student research: an integrated mixed-methods systematic review and meta-analysis. *PLoS One* 2015;10:e0127470.
 - 18 Ejaz K, Shamim MS, Shamim MS, *et al.* Involvement of medical students and fresh medical graduates of Karachi, Pakistan in research. *J Pak Med Assoc* 2011;61:115–20.
 - 19 Nicholson JA, Cleland J, Lemon J, *et al.* Why medical students choose not to carry out an intercalated BSc: a questionnaire study. *BMC Med Educ* 2010;10:25-6920-10-25.
 - 20 Acmedsci.ac.uk. INSPIRE | The Academy of Medical Sciences, 2021. Available: <<https://acmedsci.ac.uk/grants-and-schemes/mentoring-and-other-schemes/INSPIRE>> [Accessed 12 May 2021].
 - 21 Blencowe N, Glasbey J, Heywood N, *et al.* Recognising contributions to work in research Collaboratives: guidelines for standardising reporting of authorship in Collaborative research. *Int J Surg* 2018;52:355–60.
 - 22 Bhangu A, Fitzgerald JE, Koliass AG. Trainee-led research Collaboratives: a novel model for delivering multi-centre studies. *ANZ J Surg* 2014;84:902–3.
 - 23 Chapman SJ, Glasbey JCD, Khatri C, *et al.* Promoting research and audit at medical school: evaluating the educational impact of participation in a student-led national collaborative study. *BMC Med Educ* 2015;15:47.
 - 24 National Research Collaborative & Association of Surgeons in Training Collaborative Consensus Group. Recognising contributions to work in research Collaboratives: guidelines for standardising reporting of authorship in Collaborative research. *Int J Surg* 2018;52:355–60.
 - 25 STARSurg Collaborative. Prognostic model to predict postoperative acute kidney injury in patients undergoing major gastrointestinal surgery based on a national prospective observational cohort study. *BJS Open* 2018;2:400–10.
 - 26 Bandyopadhyay S, Shortland T, Wadanamby SW, *et al.* Global health education in UK medical schools (GHEMS) study protocol. *J Glob Health Rep* 2019;3:e2019052.
 - 27 European Medicines Agency. Available: https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e-6-r2-guideline-good-clinical-practice-step-5_en.pdf [Accessed 12 May 2021].
 - 28 Lombardo R, Beh EJ. Simple and multiple correspondence analysis for ordinal-scale variables using orthogonal polynomials. *J Appl Stat* 2010;37:2101–16.
 - 29 Textor J, van der Zander B, Gilthorpe MS. Robust causal inference using directed acyclic graphs: the R package 'dagitty'. *Int J Epidemiol* 2016;45:dyw341–1894.

Appendix S1

Proforma of questionnaire, to be converted to digital format for administration. Text in red describes the available answers for each question.

Q1a. Which medical school do you attend? (If currently intercalating at a separate university, please give the university from which you will receive your main degree).

Please select your university from the dropdown list
as in appendix S2

Q1b. Which year of medical school are you currently in? (Ordinal)

Please select which year of your medical degree you are currently enrolled in (First/Second/Third/ /Year 4 of 5/ Year 4 of 6/Year 5 of 6/ Year 5 of 5 /Year 6 of 6/ Intercalation/ Graduate Entry first/ Graduate Entry second/ Graduate Entry third/ Graduate Entry fourth)

Q2. Have you already completed an academic degree (Bachelor's/Masters/Doctorate)? (binary)

Please select whether you currently already have a completed higher education degree, such as a Bachelor's, Master's, or Doctorate. (YES/NO)

Q3. If you answered yes to Question 2, please select all those degrees you currently have. (Nominal/Ordinal)

Please select all degrees you currently have. (Bachelor's – in a scientific degree/Bachelor's – in an arts degree/Master's – in a scientific degree/Master's – in an arts degree/Doctorate/ Other (please type))

Q4. Choose one option that best describes your ethnic group or background? (Nominal)

Please select which ethnic group you most strongly identify as. (English, Welsh, Scottish, Northern Irish, British/Irish/Gypsy or Irish Traveller/Any other White background/ White and Black Caribbean/ White and Black African/ White and Asian/ Any other Mixed or Multiple ethnic background/ Indian/ Pakistani/ Bangladeshi/ Chinese/ Any other Asian background/ African/Caribbean/ Any other Black, African, or Caribbean background/Arab/ Any other ethnic group)

*Note, list taken from government list of ethnic groups <https://www.ethnicity-facts-figures.service.gov.uk/style-guide/ethnic-groups>

Q5. What best describes your gender (Nominal)

Female/ Male/ Prefer not to say/ Prefer to self-describe (free text)

Q6. Do you identify as LGBTQ+? (Binary)

Please select whether you identify as LGBTQ+ (Lesbian, Gay, Bisexual, Transgender, Queer/Questioning, or any other part of the LGBTQ+ community). (YES/NO)

Q7. During school, at any point, were you eligible for free school meals? (binary)

Please select whether you were, at any point, eligible for free school meals during school. (Yes/No)

Q8. How many of your first degree relatives are or have been a healthcare professional*? (Ratio/Ordinal)

Please select the number of relatives who currently or have ever identify/identified as healthcare professionals. (0,1,2,3,4,5,6,7,8,9,10,10+)

*Healthcare professional defined according to WHO international classification of healthcare workers as professionals who study, advise on or provide preventive, curative, rehabilitative and promotional health services based on an extensive body of theoretical and factual knowledge in diagnosis and treatment of disease and other health problems. This includes:

- General medical practitioners
- Specialist medical practitioners
- Nursing professionals
- Midwifery professionals
- Traditional and complementary medicine professionals
- Paramedical practitioners
- Dentists
- Pharmacists
- Environmental and occupational health and hygiene professionals
- Physiotherapists

- Dieticians and nutritionists
- Optometrists and ophthalmic opticians

Q8b. If you answered >0 Question 8, how many of your relatives are or have been held an academic position* in the healthcare environment? (Ratio/ordinal)

Please select all the number of relatives who currently or have ever identify/identified as academics in the healthcare environment. (0,1,2,3,4,5,6,7,8,9,10,10+)

*Academic position defined as working within a research or teaching capacity within a higher education institute. This includes, but is not limited, to, professors, readers, senior lecturers, lecturers, post-doctoral researchers, research fellows, teaching fellows

Q9. How many of your first degree relatives are or have been in academia? (Ratio/ordinal)

Please select the number of relatives who currently or have ever identify/identified as being in academia. (0,1,2,3,4,5,6,7,8,9,10,10+)

Q9b. If you answered >0 Question 9, how many of your relatives are or have been held an academic position in the healthcare environment? (Ratio/ordinal)

Please select all the number of relatives who currently or have ever identify/identified as academics in the healthcare environment. (0,1,2,3,4,5,6,7,8,9,10,10+)

*Academic position defined as working within a research or teaching capacity within a higher education institute. This includes, but is not limited, to, professors, readers, senior lecturers, lecturers, post-doctoral researchers, research fellows, teaching fellows

Q11. In which area did you undertake the majority of your pre-university education?

Please select the geographical area where you undertook the majority of your education before university (UK, EU, outside EU)

Q12. How well do you feel your medical school has educated you about research? (Ordinal)

Please select which best represents your opinion. (Not well at all/ Less than adequately/Adequately/More than adequately/Very well)

Q13. How much research have you undertaken to date? (Ordinal)

Please select how much research you feel you have undertaken to date. (None/A little bit/ A fair amount/Quite a lot/A very significant body of research)

Q14. If you have undertaken any research, was any of this a compulsory part of your degree? (Binary)

Please select whether any research you have done included a compulsory part of your degree. (Yes/No)

Q15. If you have not yet been involved in voluntary research, what have been the barriers preventing this? (Nominal)

Please describe the barriers that have prevented you from undertaking research outside of any compulsory degree obligation. (Free text up to 500 words)

Q16. Please tick any of the following that apply to your research. You may tick more than one box or none at all : (Nominal)

Please tick any of the following that apply to your research: (QIP/audit/basic science project/ clinical project/ co-author on original paper in peer reviewed journal/ co-author on any publications related to research/ named collaborator on original paper in peer reviewed journal/ named collaborator on any publication related to research/ presented a poster/ given an oral presentation/other (please type))

Q17. Why did/do you do research? (Nominal)

Please tick any of the following that apply to your research:
(Interest in scientific problems/Interest in the subject/Personal development/Extra income/contribution to better health care/Improving critical thinking/Career progression/Intellectual stimulation/Feel obliged to do it/Other (please type))

Q18. How much do you think research is useful in combination with your medical studies? (Ordinal)

Please select to what degree you feel research is useful in combination with your medical studies. (Not at all useful/Somewhat useful/ A little useful/A lot useful/Very useful)

Q19. How difficult is it to combine research with your medical studies? (Ordinal)

Please select how difficult you find it to integrate research with your medical studies. (Not at all difficult/Somewhat difficult/A little difficult/ A lot difficult/ Very difficult)

Q20. How much do you agree with the statement: "I wish to pursue an academic career ." (Ordinal)

Please select how strongly you intend to pursue an academic career. (Strongly disagree/ Somewhat disagree/No opinion/Somewhat agree/Strongly agree)

Q21. How much do you agree with the statement: "I wish to pursue an academic training pathway ." (Ordinal)

Please select how strongly you intend to pursue an academic career. (Strongly disagree/ Somewhat disagree/No opinion/Somewhat agree/Strongly agree)

Q22. How much do you agree with the statement: "I would be interested in undertaking (more) research in the future." (Ordinal)

Please select how strongly you agree with the above statement. (Strongly disagree/ Somewhat disagree/No opinion/Somewhat agree/Strongly agree)

**Q23. What would encourage your involvement in research in the future?
(Ordinal)**

Please tick all that apply
(More time/More incentives/ Easier access to research groups and projects/ Clearer information about how to get involved/ Clearer information about benefits of research/ Guaranteed rewards e.g. publication or presentation/ Other (please type)

Appendix S2

A list of medical schools recognised by the GMC as of 1st January 2020:

The University of Aberdeen
The University of Birmingham
The University of Bristol
The University of Buckingham
The University of Cambridge
Cardiff University
The University of Dundee
The University of East Anglia
The University of Edinburgh
University of Exeter Medical School
The University of Glasgow
The Imperial College of Science, Technology and Medicine
Keele University
King's College London
Lancaster University
The University of Leeds
The University of Leicester
The University of Liverpool
The University of London
The University of Manchester
The University of Newcastle
The University of Nottingham
The University of Oxford
Plymouth University Peninsula Schools of Medicine and Dentistry
Queen Mary University of London
The Queen's University of Belfast
St George's Hospital Medical School
Swansea University
The University of Sheffield
The University of Southampton
University College London
The University of Warwick
A combination of the University of Brighton and the University of Sussex
A combination of the University of Hull and the University of York

Appendix S3

This document outlines the participant facing information used in recruitment for the SMART study

1. **Enrolment and 1st Survey Phase:** Link shared via email and social media
 - a. Initiation: 1st – 14th November
2. **1st and 2nd Follow up for 1st Survey**
 - a. 14th November – 30th November
3. **3rd Follow up for 1st Survey**
 - a. 31st November – 1st January

1. Initial email to students

Dear Students,

We are a group of researchers from across the country who are interested in understanding the medical student perceptions of research and research-orientated careers.

We invite you to participate in a 15-minute voluntary survey that is collecting information on your views on research and research-orientated careers. When you access the survey, you are going to be asked a variety of questions about your background, your previous experiences of research, and your thoughts on a future research career. Your IP address will not be recorded. All answers will remain anonymous and confidential.

If you wish to partake in our prize draw, please provide your email address when requested by the survey. This email address will not be linked to your answers. More information on the risks and benefits of completing this survey are included in the survey link. **You can access the survey here:**

<INS LINK>

If you have any questions, you can contact a member of the research team by email at soham.bandyopadhyay@st-hildas.ox.ac.uk
Thank you for your consideration.

Many thanks,
<Your name>
SMART team

2. First and Second Follow-up email for 1st survey

Dear Students,

Just a friendly reminder that the SMART survey will remain open until January 1st, 2022. Please take a few moments to participate in this survey if you have not done so yet.

We invite you to participate in a 15-minute voluntary survey that is collecting information on your views on research and research-orientated careers. When you access the survey, you are going to be asked a variety of questions about your background, your previous experiences of research, and your thoughts on a future research career. Your IP address will not be recorded. All answers will remain anonymous and confidential.

If you wish to partake in our prize draw, please provide your email address when requested by the survey. This email address will not be linked to your answers. More information on the risks and benefits of completing this survey are included in the survey link. **You can access the survey here:**

<INS LINK>

If you have any questions, you can contact a member of the research team by email at soham.bandyopadhyay@st-hildas.ox.ac.uk

Thank you for your consideration.

Many thanks,
<Your name>
SMART team

3. Third Follow-up email for 1st survey (sent out the day before the survey closes)

Dear Students,

Just a friendly reminder that the SMART survey will remain open until January 1st, 2022. Please take a few moments to participate in this survey if you have not done so yet.

We invite you to participate in a 15-minute voluntary survey that is collecting information on your views on research and research-orientated careers. When you access the survey, you are going to be asked a variety of questions about your background, your previous experiences of research, and your thoughts on a future research career. Your IP address will not be recorded. All answers will remain anonymous and confidential.

If you wish to partake in our prize draw, please provide your email address when requested by the survey. This email address will not be linked to your answers. More information on the risks and benefits of completing this survey are included in the survey link. **You can access the survey here:**

<INS LINK>

If you have any questions, you can contact a member of the research team by email at soham.bandyopadhyay@st-hildas.ox.ac.uk

Thank you for your consideration.

Many thanks,
<Your name>

SMART team

Appendix S4

Thank you for taking part in this research study. This information is intended for anyone who may be concerned or distressed after participating in this study. If this applies to you, I would like to point out that there are several sources of advice or help which are free and readily available to you and which may prove useful. Specifically, these include (please note, you may need to copy and paste the web addresses into your browser):

Your GP: Book an appointment with your GP. They can offer advice or refer you to other more specific services to get help.

Your University Counselling Service (where available)

NHS 111

FRANK: For friendly, confidential drugs advice, call FRANK on: 0300 123 6600 (0800776600 in Scotland) . Or, you can access their website at: <http://www.talktofrank.com>

MIND: Mind is a mental health charity serving England and Wales. To ask about mental health or mental health services, or to find out more about a particular Mind service, phone the MIND Information Line on 0300 123 3393. Trained staff and volunteers will be able to give you details of services in your area and talk through options with you. <https://www.mind.org.uk/>

Students Against Depression: An award-winning website offering information, blogs and resources to help students find their way forward from low mood or depression: <http://www.studentsagainstdepression.org>

Student Minds: Student Minds is a national student mental health charity working to encourage peer interventions for student mental health. For more information, take a look at their website at: <http://www.studentminds.org.uk>

The Samaritans: Trained volunteers are able to listen to you any time day or night. We can help you talk through whatever is troubling you, find the answers that are right for you, and offer support. Call on: 116 123 <https://www.samaritans.org/>

LGBT foundation: a national charity delivering advice, support and information services to lesbian, gay, bisexual and trans (LGBT) communities <https://lgbt.foundation/>

IN AN EMERGENCY: If you are experiencing suicidal thoughts and think that you might be unable to keep yourself safe, visit your nearest A&E department or call 999.

Appendix S5

Survey of Medical student Attitudes to Research and Training pathways (SMART) study

Thank you for your interest in participating in this study. Please take a moment to read the following information, before ticking the box to confirm your participation in the study.

If you have any questions or concerns, please contact the principal researcher, Soham Bandyopadhyay, at soham.bandyopadhyay@st-hildas.ox.ac.uk

- **What is the aim of this study?** This study aims to ascertain current medical student involvement with research. We also hope to identify factors encouraging and discouraging students from partaking in research and to consider what may encourage more engagement with scientific research in the future.
- **Why have I been selected to take part?** You are being invited to take part in the questionnaire as you are a medical student currently studying for a UK medical degree at a UK medical school recognised by the General Medical Council (GMC).
- **What do I have to do?** If you choose to participate in this voluntary survey, you will be asked to complete a questionnaire about your background, your previous exposure to research and your feelings towards a research career. This study is voluntary. If you decide not to participate this will not impact your academic standing in any way. If you decide to take part, you will be asked to complete the survey by clicking on the link below. This survey is expected to take about 10 -15 minutes to complete, but there is no time limit and you can take as much time as you like. No background knowledge is required. We will ask for your consent for the collection and storage of data in accordance with the UK General Data Protection Regulation (GDPR) within the survey. For more information on GDPR please click on the following link: <https://gdpr-info.eu/>.
- **Do I have to participate?** Please note that your participation is voluntary. You may withdraw at any point during the questionnaire for any reason, before submitting your answers, by closing the browser. In cases of withdrawal from the study, no new data will be collected or linked to other data from that point on. If you do not want to answer some of the questions you do not have to, but you can still be in the study. All questions are optional. Your decision whether or not to be part of the study will not affect your academic standing or your access to university support services. If you have already submitted data and wish to withdraw from the study, please contact soham.bandyopadhyay@st-hildas.ox.ac.uk by 1st October 2021.
- **Who has approved this study?** This project has received ethics clearance through the University of Oxford's ethical approval process for research involving human participants, reference R73479/RE001.
- **How will my data be used?** Your answers will be completely anonymous, and we will take all reasonable measures to ensure that they remain confidential. Your data will be stored in a password-protected file and may be

used in academic publications. Your IP address will not be stored. If you provide us with your email address, we will delete that information at the end of the study. No answers will be linked to your email address. Research data – your anonymised answers – will be stored for a minimum of ten years after publication or public release.

- **Who will have access to my data?** Qualtrics is the data controller with respect to the personal data they hold about you and, as such, will determine how your personal data is used. Please see their privacy notice here: <https://www.qualtrics.com/privacy-statement>. Qualtrics will share any email address you provide and your anonymised answers with the University of Oxford, for the purposes of research. Researchers involved in the project will have access to this anonymised data. The University of Oxford is the data controller of university email addresses, please see their privacy notice here: <https://compliance.admin.ox.ac.uk/student-privacy-policy>. Responsible members of the University of Oxford and funders may be given access to data for monitoring and/or audit of the study to ensure we are complying with guidelines, or as otherwise required by law.
- **Are there any benefits to taking part?** Despite not have any immediate individual benefits by participating in this survey, you are given the opportunity to contribute to valuable and innovative research which could be used in the future by medical universities and the world. You may find this survey an opportunity to self-reflect. There will be the option to submit email address in order to be entered into a prize draw. At the conclusion of data collection, two random participants will be awarded £50 in Amazon vouchers, two further random participants will be awarded £25 in Amazon vouchers. This will be optional, as it requires you to provide personally identifying data (i.e. contact details). These will not be linked to the questionnaire answers given, and will only be used for contact regarding relevant rewards as above.
- **Will the research be published?** The findings of the study may be published in peer reviewed journals, presented at relevant conferences and meetings and a summary of the findings will be made available on social media.
- **Are there any possible risks involved with my participation?** Some of the questions that we ask may cause upset. If you experience any distress from participating in this study, you may stop the survey at any time or skip any upsetting questions. If your distress continues after leaving the survey, we have provided a list of supportive services nationwide that can be helpful and that you might consider contacting (appendix S5, to be linked here, and appear again at the close of the survey).
- **Who do I contact if I have a concern about the study or I wish to complain?** If you have a concern about any aspect of this project, please speak to the researcher soham.bandyopadhyay@st-hildas.ox.ac.uk who will do their best to answer your query. The researchers should acknowledge your concern within 10 working days and give you an indication of how they intend to deal with it. If you remain unhappy or wish to make a formal complaint, please contact the Chair of the Medical Sciences Inter-Divisional Research Ethics Committee: Email: ethics@medsci.ox.ac.uk; Address: Research Services, University of Oxford, Wellington Square, Oxford OX1 2JD OR The Chair will seek to resolve the matter in a reasonably expeditious manner.

- **How do I find out what was learned in this study?** This study is expected to be completed by approximately March 2022. If you would like a brief summary of the results, please write to us by email to request information
- **Who to contact for further details?** For any further questions or more information on the study, please contact us on the following email address: soham.bandyopadhyay@st-hildas.ox.ac.uk. Alternatively, you could contact principal investigator Dr Catherine Swales at catherine.swales@ndorms.ox.ac.uk.

Please note that you may only participate in this survey if you are 18 years of age or over.

I certify that I am 18 years of age or over

If you have read the information above and agree to participate with the understanding that the data (including any personal data) you submit will be processed accordingly, please check the relevant box below to get started.

Yes, I agree to take part

Tick this box if you would like to be considered for the prize draw, and are happy to be contacted via your email address for such

Yes, I would like to be entered into the draw

Appendix S6

Many thanks for filling out the SMART questionnaire. We ask you to kindly give us a further few minutes of your time to answer questions regarding the questionnaire experience in order to improve it for future users.

Approximately how many minutes did it take you to complete the questionnaire?
Numerical

In your opinion was the questionnaire *too long/ just right/ too short?*

Did you find the questions clear and easy to understand? *Yes/No*

Does the questionnaire omit any issues or factors you consider important to investigate? *Yes/No with white space*

Did you have any problems completing the questionnaire? *White space*

Is there anything else you would like us to know about the questionnaire, or the distribution of the questionnaire? *White space*

Appendix S7

We will utilise the test of parallel lines, and the output of our Multiple Correspondence Analysis to test initial statistical assumptions. We may relax the assumptions of the model by allowing partial proportional odds, or a generalised ordinal logit to ensure valid statistical analysis. We have used the directed acyclic graph approach to minimize bias introduction for causal inference from our predictive model, through selective covariate and predictor variable selection.

Our proposed initial model is: **Going to pursue academic career = (Ethnicity + Gender + Medical School + Free school meals + 1st Degree relatives in academia + Year of study + LGBTQ+)** (Equation 1).

Equation 1.

$$\text{logit}[P(\text{Going to pursue academic career} \leq j)] = \alpha_j + \beta_{\text{Ethnicity}}X_{\text{Ethnicity}} + \beta_{\text{Gender}}X_{\text{Gender}} + \beta_{\text{Medical School}}X_{\text{Medical School}} + \beta_{\text{Free School Meals}}X_{\text{Free School Meals}} + \beta_{\text{Relatives in academia}}X_{\text{Relatives in academia}} + \beta_{\text{Year of study}}X_{\text{Year of study}} + \beta_{\text{LGBTQ+}}X_{\text{LGBTQ+}}.$$

Where 'Desire to pursue academic career' represents answers to Question 20. (How much do you agree with the statement: "I wish to pursue an academic career or an academic training pathway ."). j represents the number of possible answers to Question 20 – 1. Ethnicity represents answers to Question 4. (Choose one option that best describes your ethnic group or background?) Gender represents answers to Question 5 (What gender do you identify as?). Medical School represents answers to Question 1a (Which medical school do you attend? (If currently intercalating at a separate university, please give the university from which you will receive your main degree)). Free School Meals represents answers to Question 7. (During School, at any point, were you eligible for free school meals?). Relatives in academia represents answers to Question 9. (How many of your first degree relatives are or have been in academia?). Year of study represents answers to Question 1b. (What year of medical school are you currently in?). LGBTQ+ represents answers to Question 6. (Do you identify as LGBTQ+?).

We will utilize likelihood ratios, Akaike information criterion, Schwarz criterion, -2 log likelihood, Pseudo-R² (Cox-Snell) to ascertain if this model has acceptable fit. If this full model does not, then we will optimise for a single model out of the set of possible nested models which will then be used as our subsequent model. We will utilise sampling weights identified in 1.5.

Appendix S8

Our proposed initial model is: **Research in future = (Ethnicity + Gender + Medical School + Free school meals + 1st Degree relatives in academia + Year of study + LGBTQ+)** (Equation 2).

Equation 2.

$$\text{logit}[P(\text{Desire to research in future} \leq j)] = \alpha_j + \beta_{\text{Ethnicity}}X_{\text{Ethnicity}} + \beta_{\text{Gender}}X_{\text{Gender}} + \beta_{\text{Medical School}}X_{\text{Medical School}} + \beta_{\text{Free School Meals}}X_{\text{Free School Meals}} + \beta_{\text{Relatives in academia}}X_{\text{Relatives in academia}} + \beta_{\text{Year of study}}X_{\text{Year of study}} + \beta_{\text{LGBTQ+}}X_{\text{LGBTQ+}}.$$

Where the above variables represent the same as in Appendix S7, apart from Desire to research in the future represents answers to Question 21.

Appendix S9

Going to pursue academic career = (Ethnicity + Gender + Medical School + Free school meals + Why did research + previous Degree + Year of study + LGBTQ+) (Equation 3).

Equation 3.

$$\text{logit}[P(\text{Going to pursue academic career} \leq j)] = \alpha_j + \beta_{\text{Ethnicity}}X_{\text{Ethnicity}} + \beta_{\text{Gender}}X_{\text{Gender}} + \beta_{\text{Medical School}}X_{\text{Medical School}} + \beta_{\text{Free School Meals}}X_{\text{Free School Meals}} + \beta_{\text{Why did research}}X_{\text{Why did research}} + \beta_{\text{Previous degree}}X_{\text{Previous degree}} + \beta_{\text{Year of study}}X_{\text{Year of study}} + \beta_{\text{LGBTQ+}}X_{\text{LGBTQ+}}.$$

Going to pursue academic career = (Ethnicity + Gender + Medical School + Free school meals + Barriers to voluntary research + Year of study + LGBTQ+)
(Equation 4).

Equation 4.

$$\text{logit}[P(\text{Going to pursue academic career} \leq j)] = \alpha_j + \beta_{\text{Ethnicity}}X_{\text{Ethnicity}} + \beta_{\text{Gender}}X_{\text{Gender}} + \beta_{\text{Medical School}}X_{\text{Medical School}} + \beta_{\text{Free School Meals}}X_{\text{Free School Meals}} + \beta_{\text{Barriers to voluntary research}}X_{\text{Barriers to voluntary research}} + \beta_{\text{Year of study}}X_{\text{Year of study}} + \beta_{\text{LGBTQ+}}X_{\text{LGBTQ+}}.$$

Where all previously defined variables are as defined in Appendix 7 or Appendix 8. 'Why did research' represents answers to Question 17. (Why did/do you do research?). 'Previous degree' represents answers to Question 2. (Have you already completed an academic degree?). Barriers to voluntary research represents the set of common themes identified in each answer to Question 15. (If you have not yet been involved in voluntary research, what have been the barriers preventing this?).]