BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers’ comments and the authors’ responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (http://bmjopen.bmj.com).

If you have any questions on BMJ Open’s open peer review process please email info.bmjopen@bmj.com
# Reporting of data on participant ethnicity and socioeconomic status in high-impact medical journals: A targeted literature review

| Journal: | BMJ Open |
|----------|----------|
| Manuscript ID | bmjopen-2022-064276 |
| Article Type: | Original research |
| Date Submitted by the Author: | 28-Apr-2022 |
| Complete List of Authors: | Buttery, Sara ; Imperial College London; NIHR Imperial Biomedical Research Centre<br>Philip, Keir; Imperial College London; NIHR Imperial Biomedical Research Centre<br>Alghamdi, Saeed; Imperial College London<br>Williams, Parris; Imperial College London, National Heart and Lung Institute<br>Quint, Jennifer; Imperial College London, NHLI; Imperial College London<br>Hopkinson, Nicholas; Imperial College London, National Heart and Lung Institute |
| Keywords: | STATISTICS & RESEARCH METHODS, GENERAL MEDICINE (see Internal Medicine), INTERNAL MEDICINE |
I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd (“BMJ”) its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge (“APC”) for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author’s Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.
Manuscript Title: Reporting of data on participant ethnicity and socioeconomic status in high-impact medical journals: A targeted literature review

Authors: Sara C Buttery1,2 BSc*, Keir EJ Philip1,2 MRCP*, Saeed M. Alghamdi1,2 MSc, Parris Williams1,2 MSc, Jennifer K Quint1,2 PhD**, Nicholas S Hopkinson1,2 PhD**

*Joint first author
** Joint senior author

Affiliations:

1) National Heart and Lung Institute, Imperial College London, London, United Kingdom

2) NIHR Imperial Biomedical Research Centre, London, United Kingdom

Corresponding Author: Dr Keir EJ Philip, The Muscle Laboratory, The National Heart and Lung Institute, Imperial College London, London, NW3 6HP. K.philip@imperial.ac.uk

Key words: ethnicity, race, socioeconomic status, participant characteristics, demographics

Word Count: 1716
ABSTRACT (236 words)

Objectives: To assess the frequency of reporting of ethnicity (or ‘race’) and socioeconomic status (SES) indicators in high-impact journals.

Design: Targeted literature review

Data sources and inclusion criteria: The 10 highest ranked general medical journals (Google scholar h5 index) were identified. Working backwards from 19/04/2021 in each journal, we selected the 10 most recent articles meeting inclusion/exclusion criteria, to create a sample of 100 articles. Inclusion criteria were, human research, reporting participant level data. Exclusion criteria were non-research article, animal/other non-human participant/subject; or no participant characteristics reported.

Primary and secondary outcome measures: The frequency of reporting of ethnicity (or ‘race’) and SES indicators.

Results: Of one hundred research articles included, 35 reported ethnicity and 13 SES. By contrast, 99 reported age, and 97 reported sex or gender. Among the articles not reporting ethnicity only 3 (5%) highlighted this as a limitation, and only 6 (7%) where SES data were missing. Median number of articles reporting ethnicity per journal was 2.5/10 (range 0 to 9). Only 2 journals explicitly requested reporting of ethnicity (or race), and 1 requested SES.

Conclusions: The majority of research published in high-impact medical journals does not include data on the ethnicity and socioeconomic status of participants, and this omission is rarely acknowledged as a limitation. This situation persists despite the well-established importance of this issue and ICMJE recommendations to include relevant demographic variables to ensure representative samples. Standardized explicit minimum standards are required.
Article Summary

Strengths and Limitations of this study

- This article demonstrates that reporting of these ethnicity and socioeconomic status in the highest impact medical journals remains poor despite the well-established importance of reporting these data, and guidelines promoting their inclusion in publications.

- This research demonstrates an issue of substantial public health importance, as inadequate reporting of research participant demographics has been identified as contributing to health inequalities, including racial and socioeconomic disparities in the impact of the COVID-19 pandemic.

- An important consideration is the potential that alternative approaches could be used in selecting research articles/journals which may alter results. However, our findings are consistent with historical research on this topic.
Background:

Information about the ethnicity and socioeconomic status of participants in clinical research is needed for the interpretation, generalisability and pooling of data as well as to inform discussion around health inequalities. The relevance of ethnicity and socio-economic status to health and biomedical research is well established but has been emphasised by the COVID-19 pandemic, during which specific ethnic groups and poorer individuals have been disproportionately affected. The causal pathways driving health disparities are complex and multifactorial, however under-reporting of participant characteristics has been identified as a potential contributory factor.

The International Committee of Medical Journal Editors (ICMJE) recommendations, and some journal instructions to authors promote inclusion of these data. Previous studies have identified that reporting is frequently incomplete with limited progress made over the last three decades. Recent years have seen an increased focus on ethnicity and socioeconomic status in medicine, however there is a lack of research as to whether this has resulted in better reporting.

To evaluate the current situation in this area, we assessed the frequency of reporting of ethnicity (or ‘race’) and socioeconomic status indicators in a sample of research articles published in high impact general medical journals in Spring 2021.

Methods:

We identified the 10 highest ranked journals as per Google scholar ‘Health and Medical (general)’ category up to April 2021. At the time of data collection these were The New England Journal of Medicine (NEJM), The Lancet, the Journal of the American Medical Association (JAMA), Proceedings of the National Academy of Sciences of the United States of America (PNAS), Nature Medicine, Public Library of Science One (PLOS One), The British Medical Journal (BMJ), Cochrane, Cell Metabolism, and Science Translational Medicine. PNAS and PLOS One include a wide range of subject areas therefore the subsections ‘Biological Sciences, Medical Science’ and ‘Clinical Medicine’ were used.
respectively. From each of these 10 journals, using the journals own websites, we worked backwards from April 19th 2021, selecting the 10 most recent journal articles that met inclusion/exclusion criteria. Inclusion criteria were: research articles, reporting participant level data. Articles were excluded if they were: not research (e.g. editorial, news, images etc.), animal/other non-human participant/subject; or no participant characteristics reported. Laboratory studies using human derived tissues or cells were included if donor information was provided. Journal reporting guidance and requirements were also assessed by evaluating author guidelines, websites, and contacting the respective editorial/publishing teams. Data were collected on which participant level characteristics were reported and how. Data were also collected on if the absence of reporting these variables was noted as a limitation. The journals’ accessible policies and guidance on reporting these variables was also reviewed. Data collection and analysis was conducted by SCB, KEJP, SMA and PW. All journals were reviewed and articles selected by at least two researchers independently, who then came together to discuss any inconsistencies with a third researcher.

Ethnicity and race are related yet different constructs and arguably the latter term should be abandoned. However, given the frequent lack of standardisation in the literature and that the terms are in practice often used interchangeably we accepted the use of either term. Similarly, regarding reporting of socioeconomic status indicators, various often inconsistent methods are used, therefore we opted to assess both direct measures such as the Index of Multiple Deprivation, but also measures from which socioeconomic status could be inferred such as educational attainment and job role. The focus being if, rather than how, such measures are reported.

Patient and Public involvement

No patients or members of the public were specifically involved in the conduct or reporting of this study. However, we consider the topic of interest to be a public health issue of clear and substantial importance.
Results

650 publications were assessed to identify 100 meeting inclusion criteria (see figure 1 and Supplementary Material). Of one hundred research articles included, 35 reported ethnicity (or race) and 13 reported socioeconomic status. By contrast, 99 reported age, and 97 reported sex or gender (Table1). Among the articles not reporting ethnicity only 3 (5%) highlighted this as a limitation, and only 6 (7%) highlighted where socioeconomic status data were missing. Median number of articles reporting ethnicity per journal was 2.5/10 (range 0/10 (PLOS One), to 9/10 (JAMA)). Only 2 journals explicitly requested reporting of participant ethnicity (or race), and 1 requested socioeconomic status. Types of research included – interventional studies (n=30), cohort studies (n=35), case-control studies (n=3), systematic reviews and metanalyses (n=16), epidemiological and surveys (n=3), and other (n=13). Twenty of the 100 were laboratory studies (either observational or involving interventional manipulation of samples) using human samples, of which 4 reported ethnicities of sample donors (of others, none mentioned as a limitation), and none reported socioeconomic status.

Among the 24 papers describing clinical trials, 50% reported ethnicity, with none highlighting the absence of these data as a limitation. 12.5% of trials reported an indicator of socioeconomic status, with one of the 21 not reporting socioeconomic status highlighting this absence as a limitation.

Of note, two of the research articles included in our sample identified ethnicity as being relevant to their research topic, yet did not provide relevant data on their study participants or highlight the lack of this data as a limitation of their study ‘in the case of DNA-based mutation testing, poor sensitivity in detecting mutations in infants from ethnic and racial minority groups’15, and ‘peripheral oxygen saturation can substantially differ from the SaO\textsubscript{2} under certain conditions and may be less accurate in Black patients than in White patients.’16.

Figure 1: Flow diagram of included/excluded articles
Table 1: Reporting of ethnicity and/or race, and Socioeconomic Status indicators in research articles

| Report participant level characteristics | N | Additional notes |
|-----------------------------------------|---|-----------------|
| Report ethnicity and/or race            | 100 |                |
| Report ethnicity and/or race            | 35/100 report | Range per journal: JAMA 9/10, with clear guidance that this information is expected. |
| Noted in limitations                    | 65 Not report | Some studies identify race and ethnicity as being relevant to the research focus, yet did not provide relevant data on their study participants or highlight this as a limitation of their study e.g. |
|                                          | 62 of the 65 do not state this as a limitation | |
|                                          | 3 Do highlight this as a limitation. | |
| Report socioeconomic status indicator   | 13/100 report at a measure of SES (6 direct measure e.g. Index of Multiple Deprivation, Poverty income ratio; 7 measures from which SES can be inferred e.g. educational attainment, job role) | 87/100 did not report any indication of SES |
| Noted in limitations                    | 6/87 identified this as a limitation | |
| Age reported                             | 99/100 | |
| Sex or Gender reported                  | 97/100 | |

Percentages not given as most results have 100 as the denominator.
Discussion

The majority of research published in high-impact medical journals does not include data on the ethnicity and socioeconomic status of participants, and this omission is rarely acknowledged as a limitation. This finding echoes related historical research,8-13 but its persistence is of concern and is surprising given current awareness of such issues17 18.

These findings have important implications for the interpretation and application of research findings, both within academia and beyond, with the ongoing omission no longer justifiable as simple oversight. As highlighted by Baker et al.19 in relation to data relating to LGBTQI+ communities, but equally relevant here, ‘Data are fundamentally political: decisions about which data are collected and which are overlooked both reflect and shape policy and program priorities.’

Our results could have multiple contributory factors. For some research including secondary data analyses, ethnicity and socioeconomic status data may not have been available to the researchers, but given the lack of explanation, it remains unclear if these data were unavailable, or available but not included in publications. The low level of reporting in controlled clinical trials suggests issues beyond unavailability of data, as in these studies such data would be simple to collect. Additionally, given research successfully reporting these data, the justification for these omissions remains unexplained.

The increased frequency of reporting ethnicity compared to socioeconomic status, may indicate differences between the perceived relevance of these variables. This would be in keeping journal author guidelines and ICMJE recommendations that encourage the inclusion of relevant demographic variables to ensure representative samples5, more often explicitly stating race and/or ethnicity, than socioeconomic status. The relevance of these factors may not have been apparent to authors and editorial teams, however ICMJE Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly work in Medical Journals5 states 'Because the relevance of such variables as age, sex, or ethnicity is not always known at the time of study design, researchers should aim for inclusion of representative populations into all study types and at a minimum provide descriptive data for these and other relevant demographic variables.’. Of note, not all of the journals in our sample state
that they follow the ICMJE recommendations. However, whether or not the journal states they follow guidance or not, this has no impact upon the relevance of these data and the importance of reporting them. Additionally, Maduka et al. found no difference between journals stating they follow ICMJE recommendations, and those that do not, in the frequency of reporting race and ethnicity in a sample of surgical research publications in 2019.

Certain considerations and limitations require highlighting. Firstly, different approaches to selecting research papers may alter findings. Secondly, we identified high-impact journals using the Google Scholar h5 index but acknowledge various other equally valid methods exist. Thirdly our analysis focused on if ethnicity and/or race was reported, but we acknowledge that these are not synonymous terms. In addition to if these variables are reported, how they are reported is also an important area for discussion and research. The widespread omissions identified by this research suggests a structural problem. Indeed, we the authors have published research which would have met the inclusion criteria and failed to report these specific characteristics. Our intention is to highlight an issue and suggest approaches to address it.

Given that inadequate reporting persists despite research highlighting the issue, author and ICMJE recommendations, and the current socio-political climate, there is a clear need for more explicit requirements that are adhered to in practice. This is likely best achieved if steps are integrated into each stage of the research process, from protocol to publication. For example, Fain et al. compared reporting of race and ethnicity on ClinicalTrials.gov before and after the requirement to report these data (if collected), was introduced, finding that this was associated with an increase from 42% to 92%. Similar explicit requirements could be taken in EQUATOR guidelines, and research ethics applications. From our sample, the journal JAMA had the most explicit guidance for reporting race and ethnicity, and this variable was reported in 9/10 of the articles we reviewed. Of note from 2022 the New England Journal of Medicine will be requiring authors of research articles to provide data on the representativeness of the sample including race or ethnic group, though it is unclear if socioeconomic status indicators will also be required.
Conclusion

The reporting of ethnicity and socioeconomic status in high-impact medical research remains poor, despite a consensus on its importance. Omission of these participant characteristics limits the interpretation, generalisability, and pooling of data, that are required to facilitate informed discussion around health inequalities. Guidance and encouragement have so far proven insufficient to change practice in this area. Standardised, explicit, minimum standards are required.

Author Contributions:

SCB, had the original idea for the study. SCB, KEJP, SMA and PW collected the data. All authors (SCB, KEJP, SMA, PW, JKQ and NSH) contributed to the design of the study. KEJP analysed the data initially, which was verified by SCB, SMA and PW. KEJP wrote the first draft of the manuscript. All authors (SCB, KEJP, SMA, PW, JKQ and NSH) critically appraised the manuscript and approved it for submission, and had full access to the data and can take responsibility for the integrity of the data and the accuracy of the data analysis. The corresponding author attests that all listed authors (SCB, KEJP, SMA, PW, JKQ and NSH) meet authorship criteria and that no others meeting the criteria have been omitted.

Funding:

KEJP is supported was supported by the Imperial College Clinician Investigator Scholarship (internal award with no specific grant number/code). The funders had no say in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Competing interests:

None reported.
Acknowledgements:

None.

Data sharing:

All data used in this study are publicly available.

Ethics Approval

Ethics approval for this study was not required as all data used are freely available in the published literature.

References

1. Williamson EJ, Walker AJ, Bhaskaran K, et al. Factors associated with COVID-19-related death using OpenSAFELY. *Nature* 2020;584(7821):430-36. doi: 10.1038/s41586-020-2521-4 [published Online First: 2020/07/09]

2. Chastain DB, Osae SP, Henao-Martinez AF, et al. Racial Disproportionality in Covid Clinical Trials. *The New England journal of medicine* 2020;383(9):e59. doi: 10.1056/NEJMp2019711 [published Online First: 2020/08/12]

3. Krieger N, Waterman PD, Chen JT, et al. Missing again: US racial and ethnic data for COVID-19 vaccination. *Lancet* 2021;397(10281):1259-60. doi: 10.1016/S0140-6736(21)00465-7 [published Online First: 2021/03/21]

4. Webb Hooper M, Napoles AM, Perez-Stable EJ. COVID-19 and Racial/Ethnic Disparities. *Jama* 2020;323(24):e2466-67. doi: 10.1001/jama.2020.8598 [published Online First: 2020/05/12]

5. ICMJE. Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly work in Medical Journals. Updated December 2019, 2019.

6. The Lancet. Information for Authors 2021 [cited 2021 02/09]. Available from: https://marlin-prod.literatumonline.com/pb-assets/Lancet/authors/tl-info-for-authors.pdf.

7. JAMA. Instructions for Authors 2021 [cited 2021 02/09]. Available from: https://jamanetwork.com/journals/jama/pages/instructions-for-authors.

8. Geller SE, Koch A, Pellettier B, et al. Inclusion, analysis, and reporting of sex and race/ethnicity in clinical trials: have we made progress? *J Womens Health (Larchmt)* 2011;20(3):315-20. doi: 10.1089/jwh.2010.2469 [published Online First: 2011/03/01]

9. Brahman D, Bauchner H. Changes in reporting of race/ethnicity, socioeconomic status, gender, and age over 10 years. *Pediatrics* 2005;115(2):e163-6. doi: 10.1542/peds.2004-1437 [published Online First: 2005/02/03]

10. Kanakamedala P, Haga SB. Characterization of clinical study populations by race and ethnicity in biomedical literature. *Etnh Dis* 2012;22(1):96-101. [published Online First: 2012/07/11]

11. Bokor-Billmann T, Langan EA, Billmann F. The reporting of race and/or ethnicity in the medical literature: a retrospective bibliometric analysis confirmed room for improvement. *Journal of clinical epidemiology* 2020;119:1-6. doi: 10.1016/j.jclinepi.2019.11.005 [published Online First: 2019/11/13]

12. Lee SJ, Kavanaugh A. A need for greater reporting of socioeconomic status and race in clinical trials. *Ann Rheum Dis* 2004;63(12):1700-1. doi: 10.1136/ard.2003.019588 [published Online First: 2004/11/18]
13. Ma IW, Khan NA, Kang A, et al. Systematic review identified suboptimal reporting and use of race/ethnicity in general medical journals. *Journal of clinical epidemiology* 2007;60(6):572-8. doi: 10.1016/j.jclinepi.2006.11.009 [published Online First: 2007/05/12]

14. Ioannidis JPA, Powe NR, Yancy C. Recalibrating the Use of Race in Medical Research. *Jama* 2021;325(7):623-24. doi: 10.1001/jama.2021.0003 [published Online First: 2021/01/26]

15. Ray TR, Ivanovic M, Curtis PM, et al. Soft, skin-interfaced sweat stickers for cystic fibrosis diagnosis and management. *Sci Transl Med* 2021;13(587) doi: 10.1126/scitransmed.abd8109 [published Online First: 2021/04/02]

16. Schjorring OL, Klitgaard TL, Perner A, et al. Lower or Higher Oxygenation Targets for Acute Hypoxic Respiratory Failure. *The New England journal of medicine* 2021;384(14):1301-11. doi: 10.1056/NEJMoa2032510 [published Online First: 2021/01/21]

17. Flanagin A, Frey T, Christiansen SL, et al. The Reporting of Race and Ethnicity in Medical and Science Journals: Comments Invited. *Jama* 2021;325(11):1049-52. doi: 10.1001/jama.2021.2104 [published Online First: 2021/02/23]

18. The New England Journal of Medicine. Race and Medicine 2021 [cited 2021 05/09/2021].

19. Baker KE, Streed CG, Jr., Durso LE. Ensuring That LGBTQI+ People Count - Collecting Data on Sexual Orientation, Gender Identity, and Intersex Status. *The New England journal of medicine* 2021;384(13):1184-86. doi: 10.1056/NEJMp2032447 [published Online First: 2021/04/02]

20. International Committee of Medical Journal Editors. Journals stating that they follow the ICMJE Recommendations 2021 [cited 2021 09/09/2021]. Available from: [http://www.icmje.org/journals-following-the.icmje-recommendations/](http://www.icmje.org/journals-following-the.icmje-recommendations/)

21. Maduka RC, Broderick M, White EM, et al. The Reporting of Race and Ethnicity in Surgery Literature. *JAMA Surg* 2021 doi: 10.1001/jamasurg.2021.3752 [published Online First: 2021/08/19]

22. Fain KM, Nelson JT, Tse T, et al. Race and ethnicity reporting for clinical trials in ClinicalTrials.gov and publications. *Contemp Clin Trials* 2021;101:106237. doi: 10.1016/j.cct.2020.106237 [published Online First: 2020/12/09]

23. EQUATOR network. Reporting guidelines 2021 [Available from: [https://www.equator-network.org/reporting-guidelines/](https://www.equator-network.org/reporting-guidelines/)] accessed 04/10/2021 2021.

24. Rubin E. Striving for Diversity in Research Studies. *The New England journal of medicine* 2021 doi: 10.1056/NEJMe2114651 [published Online First: 2021/09/14]

**Figure legend**

Figure 1: Flow diagram of study inclusion/exclusion
Figure 1: Flow diagram of study inclusion/exclusion

Journals with the highest Google Scholar h5 index: NEJM, The Lancet, JAMA, PNAS, Nature Medicine, PLOS One, BMJ, Cochrane, Cell Metabolism, and Science Translational Medicine

Working backwards from April 19th 2021 through each journal until 10 research articles included meeting inclusion criteria. 650 publications assessed in total to reach target number of articles.

Excluded (n=550)
- Non-research (n=442)
- Animal/other non-human participant/subject (n=83)
- No participant characteristics reported (n=25)

100 research publications included (x10 per journal)
- 88 Full articles
- 12 Research Letters/Brief Communications (All additional supplementary data also reviewed)
| Journal          | Specialty focus | Date of pub        | Title                                                                 | DOI                           | Disease focus        |
|------------------|-----------------|--------------------|----------------------------------------------------------------------|-------------------------------|----------------------|
| NEJM             | Gastroenterology | arApril 15 2022    | Hypothermic DOI: 10.1053/jgastro.2022.07.054                        |                               | Liver transplantation |
| NEJM             | Psychiatry      | arApril 15 2022    | Trial of PsilocinDOI: 10.1053/jpsych.2022.07.054                     |                               | Depression           |
| NEJM             | Virology/Vaccinology | arApril 15 2022 | BNT162b2 m DOI: 10.1053/jvaccin.2022.07.054                         |                               | COVID-19             |
| NEJM             | Intensive care  | arApril 15 2022    | Dexamethadon DOI: 10.1053/jic.2022.07.054                           |                               | Sepsis               |
| NEJM             | Renal           | arApril 15 2022    | Lenalidomide DOI: 10.1053/jrenal.2022.07.054                        |                               | Renal cancer         |
| NEJM             | Intensive care  | arApril 15 2022    | Lower or Higher DOI: 10.1053/jic.2022.07.054                         |                               | Renal cancer         |
| NEJM             | Endocrinology   | arApril 15 2022    | Glycemic Indices DOI: 10.1053/jendocr.2022.07.054                    |                               | Cardiovascular       |
| NEJM             | Haematology     | arApril 15 2022    | Sutimlimab DOI: 10.1053/jhaemat.2022.07.054                         |                               | Cold agglutination   |
| NEJM             | Virology/Vaccinology | arApril 15 2022 | Antibody Response DOI: 10.1053/jvaccin.2022.07.054                   |                               | SARS-CoV-2           |
| NEJM             | Gastroenterology | arApril 15 2022    | Adjuvant Niv DOI: 10.1053/jgastro.2022.07.054                        |                               | GI malignancy        |
| The Lancet       | Virology/Vaccinology | arApril 15 2022 | Thromboembolism DOI: 10.1053/jthromb.2022.07.054                     |                               | COVID-19             |
| The Lancet       | Infection       | arApril 15 2022    | VI Effect of infusion DOI: 10.1053/jinfec.2022.07.054                 |                               | SARS-CoV-2           |
| The Lancet       | Virology/Vaccinology | arApril 15 2022 | Vol1 SARS-CoV-2 DOI: 10.1053/jvaccin.2022.07.054                     |                               | SARS-CoV-2           |
| The Lancet       | Virology/Vaccinology | arApril 15 2022 | Efficacy of Ch DOI: 10.1053/jvaccin.2022.07.054                      |                               | COVID vaccination     |
| The Lancet       | Neurology       | arApril 15 2022    | The SANAD 1 DOI: 10.1053/jneuro.2022.07.054                         |                               | Epilepsy             |
| The Lancet       | Neurology       | arApril 15 2022    | The SANAD 2 DOI: 10.1053/jneuro.2022.07.054                         |                               | Epilepsy             |
| The Lancet       | Virology/Vaccinology | arApril 15 2022 | Virus Vaccine DOI: 10.1053/jvaccin.2022.07.054                      |                               | HIV (in press)       |
| The Lancet       | Renal medicine  | arApril 15 2022    | Comparison of DOI: 10.1053/jrenal.2022.07.054                        |                               | Renal replacement     |
| The Lancet       | Obstetrics      | arMarch 27, 2022   | Trends in Prematurity DOI: 10.1053/jobst.2022.07.054                  |                               | Pneumonia            |
| JAMA             | Endocrinology   | arApril 15 2022    | Effect of Sub DOI: 10.1053/jendocr.2022.07.054                      |                               | Obesity              |
| JAMA             | Endocrinology   | arApril 15 2022    | Effect of Con DOI: 10.1053/jendocr.2022.07.054                      |                               | Obesity              |
| JAMA             | Virology/Vaccinology | arApril 15 2022 | Effect of Iver DOI: 10.1053/jvaccin.2022.07.054                     |                               | COVID-19             |
| JAMA             | Virology/Vaccinology | arApril 15 2022 | Binding and DOI: 10.1053/jvaccin.2022.07.054                      |                               | COVID-19             |
| JAMA             | Intensive care  | arApril 15 2022    | Discriminant DOI: 10.1053/jic.2022.07.054                           |                               | COVID-19             |
| JAMA             | Cardiovascular  | arApril 15 2022    | Effect of Low DOI: 10.1053/jcardiovasc.2022.07.054                  |                               | Peripheral vascular   |
| JAMA             | Oncology        | arApril 15 2022    | Effect of Ceido DOI: 10.1053/jonc.2022.07.054                        |                               | Colon cancer         |
| JAMA             | Microbiology    | arApril 15 2022    | Antibacterial DOI: 10.1053/jmicrobiol.2022.07.054                   |                               | Antibiotics          |
| JAMA             | Obstetrics      | arApril 15 2022    | Trends in Age DOI: 10.1053/jobst.2022.07.054                         |                               | Menopause            |
| JAMA             | Intensive care  | arMarch 23, 2022   | Intubation Prd DOI: 10.1053/jic.2022.07.054                          |                               | Critical illness      |
| PNAS             | Oncology        | arMarch 30, 2022   | Estrogen rece DOI: 10.1053/jonc.2022.07.054                         |                               | Prostate Cancer      |
| PNAS             | Virology/Vaccinology | arMarch 30, 2022 | Health and DOI: 10.1053/jvaccin.2022.07.054                        |                               | AMR Resist           |
| PNAS             | Haematology     | ar16th March 2022  | Loss of expression DOI: 10.1053/jhaemat.2022.07.054                 |                               | Leukemia             |
| PNAS             | Virology/Vaccinology | ar9th March 2022 | Influence of DOI: 10.1053/jvaccin.2022.07.054                       |                               | COVID-19             |
| PNAS             | Infectious care | ar9th March 2022   | Elevated Cere DOI: 10.1053/jinfectious.2022.07.054                  |                               | TB                   |
| PNAS             | Endocrinology   | ar2nd March 2022   | Glucagon blo DOI: 10.1053/jendocr.2022.07.054                        |                               | T2DM                 |
| PNAS             | Virology/Vaccinology | ar23rd Feb 2022 | Modeling SAI DOI: 10.1053/jvaccin.2022.07.054                      |                               | COVID-19             |
| PNAS             | Oncology        | ar9th Feb 2022     | Arsenic trioxide DOI: 10.1053/jonc.2022.07.054                      |                               | Leukemia             |
| PNAS             | Oncology        | ar2nd Feb 2022     | Efficient detection DOI: 10.1053/jonc.2022.07.054                    |                               | Colon Cancer         |
| PNAS             | Virology/Vaccinology | 5th Jan 2021 | Data-driver DOI: 10.1053/jvaccin.2022.07.054                       |                               | COX2 inhibitor        |
| Nature Medicine  | Respiratory     | arVolume 27 2022   | Integrative mhttps://www.Bronchiectat:https://www.Ai |                               |                      |
| Nature Medicine  | Digital medicine | arVolume 27 2022 | Assessment https://www.Ai |                               |                      |
| Nature Medicine  | Infectious care | arVolume 27 2022   | Malaria is a https://www.Malaria/iro |                               |                      |
| Nature Medicine  | Virology/Vaccinology | arVolume 27 2022 | Attributes ahttps://www.long COVID |                               |                      |
| Nature Medicine  | Infectious care | arVolume 27 2022   | Developmenthttps://www.dermatitis |                               |                      |
| Nature Medicine  | Reproduction   | arVolume 27 2022   | Fetal growth https://www.Fetal growt |                               |                      |
| Nature Medicine  | Haematology    | arVolume 27 2022   | Altered pervwww.natur ALS |                               |                      |
| Nature Medicine  | Haematology    | arVolume 27 2022   | Homozygous www.multiple m |                               |                      |
| Nature Medicine  | Neurology      | arVolume 27 2022   | Impaired mehttps://www.Meningeal |                               |                      |
| Nature Medicine  | Oncology       | arVolume 27 2022   | TCR-engineer https://www.Cancer |                               |                      |
| PLOS One (clinical medicine) | Urology | continuous April 19, 2022 Effect of diethttps://doi.Uroinary calk |                               |                      |
NEJM Non-research articles (n=24); Animal studies/Other non human (n=0); human research with no participant level data reported (n=2)
Lancet Non-research articles (n=82); Animal studies/Other non human (n=0); human research with no participant level data reported (n=2)
JAMA Non-research articles (n=51); Animal studies/Other non human (n=0); No participant level data reported (n=2)
PNAS Non-research articles (n=69); Animal studies/Other non human (n=30); No participant level data reported (n=4)
NatureMed: Non-research articles (total n=18); Animal studies/Other non human (n=1); No participant level data reported (n=7)
PLOSOne Non-research articles (n=0); Animal studies/Other non human (n=4); No participant level data reported (n=0)
BMJ Non-research articles (n=141); Animal studies/Other non human (n=1); No participant level data reported (n=2)
Cochrane Non-research papers (n=0); Animal studies/Other non human (n=0); No participant level data reported (n=1, but only because no studies included)
Cell metabolism Non-research articles (n=54); Animal studies/Other non human (n=33); No participant level data reported (n=0)
SciTransMed Non-research articles (n=3); Animal studies/Other non human (n=14); No participant level data reported (n=5)
| Country | Research Type | Study Design | Data | Additional Details |
|---------|---------------|--------------|------|--------------------|
| USA/UK  | Original research (full paper) | Cohort study | Yes: age, gender (M/F) | |
| USA/UK  | Original research (full paper) | Cohort study | Yes: sex (M/F) | |
| USA/UK  | Original research (full paper) | Cohort study | Yes: Age, mNo (and not) | |
| USA/UK  | Original research (full paper) | Cohort study | Yes: Gende No (and not) | |
| USA/UK  | Original research (full paper) | Cohort study | Yes: Sex (bcNo) (and not) | |
| USA/UK  | Original research (full paper) | Cohort study | Yes: 8 fem:No (and not) | |
| USA/UK  | Original research (full paper) | Systematic review (no meta-analysis) | Yes; age, all women | |
| USA/UK  | Original research (full paper) | Cohort study | Yes: age, (aNo) (and not) | |
| US      | Original research (full paper) | Cohort study | Yes; mean :yes; white e | |
| UK      | Original research (full paper) | Cohort study | Yes; sex, ag:yes, white, i | |
| UK      | Original research (full paper) | Cohort study | Yes; gender:yes; ethnic | |
| UK      | Canada Original research (full paper) | Cohort study | Yes, age, se:yes | |
| UK      | China Original research (full paper) | Systematic review and meta-analysis | Yes; age, %No (and not) | |
| UK      | China Austr Original research (full paper) | Systematic review and m:Yes: Age, 's:No (and not) | |
| UK      | Original research (full paper) | Systematic review and m:Yes: Age, if:No (and not) | |
| UK      | Original research (full paper) | Systematic review and m:Yes: Age, (fNo) (and not) | |
| UK      | Original research (full paper) | Systematic review and m:Yes: Age, %No (and not) | |
| UK      | Original research (full paper) | Systematic review and m:Yes: age, geNo (and not) | |
| UK      | Original research (full paper) | Systematic review and m:Yes: age, %No (and not) | |
| UK      | Original research (full paper) | Systematic review and m:Yes: age, alNo (and not) | |
| UK      | Original research (full paper) | Systematic review and m:Yes: age, seYes - as per | |
| UK      | China Original research (full paper) | Cohort study | Yes, Age, se:No (and not) | |
| UK      | USA Original research (full paper) | Other Observational (lab Yes, Age, se:No (and not) | |
| UK      | Scotland Original research (full paper) | Other Observational (lab Yes, Age, se:No (and not) | |
| UK      | Original research (full paper) | Other Observational (lab Yes, age, se:No (and not) | |
| UK      | USA Original research (full paper) | Other Observational (lab Yes, age, se:No (and not) | |
| UK      | USA Original research (full paper) | Other Observational (lab Yes, age, se:No (and not) | |
| UK      | USA Original research (full paper) | Other Observational (lab Yes, age, se:No (and not) | |
| UK      | USA Original research (full paper) | Other Observational (lab Yes, age, se:No (and not) | |
| UK      | USA Original research (full paper) | Other Observational (lab Yes, age, se:No (and not) | |
| UK      | Switzerland Original research (full paper) | Other Observational (lab Yes, age, se:No (and not) | |
| USA     | USA Original research (full paper) | Other Observational (lab Yes, age, se:No (and not) | |
| USA     | France Original research (full paper) | Other Observational (lab Yes, Age, gtNo (and not) | |
| USA     | USA Original research (full paper) | Cohort study | Yes male to:No (and not) | |
| USA     | Hong Kong Original research (full paper) | Interventional other (LabYes : sex m:No (and not) | |
| USA     | Switzerland Original research (full paper) | Interventional other (LabYes (under No) (and not) | |
| USA     | USA Original research (full paper) | Interventional other (LabYes Age, ge No. (and no | |
| USA     | USA Original research (full paper) | Cohort study | Yes (supp t:Yes: race ‘A’ | |
| USA     | Spain Original research (full paper) | Cohort study | Yes: Age, gtNo (and not) | |
| USA     | UK Original research (full paper) | Cohort study | Yes: Age, re:Yes: race (W | |
| USA     | USA Original research (full paper) | Interventional other (LabYes: age, ge:Yes White E | |

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml
participant level data reported (n=2)
no participant level data reported (n=2)
reported (n=2)
ported (n=4)
vel data reported (n=7)
ported (n=0)
orted (n=2)
ported (n=1, but only because no studies included)
vel data reported (n=0)
ata reported (n=5)
Yes: ethnicity; not hispanic or latino/ hispanic or latine; race; white, asian, black or african American, more than one race, unknown or not reported.

Yes: 'population sector' general jewish, Arab, Ultra-orthodox jewish.

Yes: %white (patient reported).

Yes: index of multiple deprivation.

Yes but Not specifically, but do provide data on Health Insurance (private/semiprivate, government subsidised, uninsured), number of people in the household.

Yes: Educational attainment, poverty (Poverty income ratio).

Yes: University level education, and employment status (employed/student/unemployed).

Yes: the inclusion of many different populations could limit uniformity (white, rYes: Index of multiple No, but has Not required!)

Yes; sex female/male

Yes: race (white, asian, black, other, not reported)

Yes: ethnicity (white, mixed race, asian, black, chinese, other, prefer not to say)

Yes: race (white, asian, black, other, not reported)

Yes: ethnicity (White, Black, Asian, Other), Hispanic ethnicity.

Yes: race/ethnicity (Non-hispanic white, hispanic, Native American, Black).

Yes: race/ethnicity (Black or Black American, White, Asian).

Yes: mixed race, Black or African American, Colombian native.

Yes: 'gender distribution' male (n).

Sex female/male

Just sex male/female

Just sex male/female

Just sex male/female

Just sex female/male

N/a

Just sex

No just %male

No

Just sex

Just sex

No

Just sex

Sex female/male

Just sex

Just sex

Yes: the inclusion of many different populations could limit uniformity (white, rYes: Index of multiple No, but has Not required!)

Yes; sex; male/female

Yes: sex female/male

Yes: sex; no of males and females

No as all male; prostate Ca.

Just women %

Just sex men/women

Just Sex, male/female. Interesting to note they use male female and men women.

Just sex male/female

Just sex male/female

Just sex male/female

Just sex

Just sex % female

No just %male

Just sex, though ?reporting gender would have been more appropriate as reported

Report Sex and Gender. If so how?

% boys

No

Just sex

No

Just sex

No

Just sex

No

Just sex

Not required!

Just sex

No

No just %male

No

Yes: the inclusion of many different populations could limit uniformity; not hispanic (male% yes; 1) App users were disproportionately female, and those over 70 years old were underrepresented, 2) Caution is needed when interpreting results in subgroups, 3) unable to analyze the impact of ethnicity due to incomplete data.

LIMITATION: index of multiple Yes (male% yes; 1) App users were disproportionately female, and those over 70 years old were underrepresented, 2) Caution is needed when interpreting results in subgroups, 3) unable to analyze the impact of ethnicity due to incomplete data.

LIMITATION: Yes 'gender' Yes 'Secondly, the impact of dietary factors on stone recurrence variety...
LIMITATION
No (and not in LinMale Vs Fe; Yes 'most studies come from East Asia, and there may be ethnic differ-
t in limitatioNo (IN LIMITATIO) No Not patient level but did state 'We also limited our search to high inc-
t in limitatioNo (and not in lin Just sex No

t in limitatioNo (and not in lin states data No
t in limitatioNo (and not in lin gender mal No
t in limitatioNo (and not in linJust sex (bc No
t in limitatioNo (and not in lin just numbe No
t in limitatioNo (and not in lin No No
t in limitatioNo (and not in lin n/a No
ethnicity or YES; as stated; yes; men n no
European no (IN LIMITATIO yes no; With respect to external validity, the trial included women from
mixed, asiano (IN LIMITATIOyes; men, wyes; it is discussed but not reported; as the analyses are unadjusted i-
t in limitatioYES education (wyes; female; some note on socio economic status but no mention of ethnicity/rac
ciety; white b YES household myes; male/f For ethnicity, we grouped together black people, Asian people, and t
yes; as stated; yes n/a
in limitatioNo (and not in linyes % wom no
ed No (and not in lin/a all wom no
ethnicity n %no (IN LIMITATIOyes; male s; socio economic status were not observed in the data set
yes; as stated; yes n/a
t in limitatioNo (and not in linyes: individ no
t in limitatioNo (and not in lin No No
t in limitatioNo (and not in lin No No
t in limitatioNo (and not in linSex reporter No
t in limitatioNo (and not in lin %female No
t in limitatioNo (and not in lin gender mal No
t in limitatioNo (and not in linAll female, no
t in limitatioNo (and not in lin and % fer no
t in limitatioNo (and not in linYes (Sex n f no
t in limitatioNo (and not in linYes (Gender no
t in limitatioNo (and not in linYes Male, F no
t in limitatioNo (and not in linYes, (doesn no
ian, black, hNo (and not in linYes (Sex n/H no
ian, african No (and not in linYes, (sex n/H no
t in limitatioNo (and not in linYes, (sex n/H no
t in limitatioNo (and not in linYes (Doesnt no
ian, Black, No (and not in linYes: sex maNo
t in limitatioNo (and not in linYes gender No
White, Black No (and not in linYes sex maNo
Black No (and not in lin No No
Pulse oximetry may be less accurate in black people. Though having not reported race-ethnicity it is unclear conclusions; however, the diversity also increases the range of differences that may be helpful in establ

Also includes marital status and vote in 2016 presidential election

0 years of age were underrepresented, 2). Caution is needed in the interpretation of associations found i

ies from age, gender, race, and region remained unknown due to the lack of related studies.  

References that restrict the generalization and reliability of the results.'

Some notes on socio-economic status but no mention of ethnicity/race; Fourthly, the community dwelling older people who participated in our study were highly educated, people of mixed ethnicity. This might have obscured differences between ethnic groups.
if their study participants are all white, all black etc. 

lishing associations, and when effects are seen, they are likely to be robust and meaningful.

in smaller population subgroups 3) unable to analyze the impact of ethnicity due to incomplete data.
 level of each record within each dataset is needed to provide assurance of high linkage quality and to a
had a high percentage of computer ownership, and lived in more affluent areas of Sydney; our results m
allow assessment of whether this varies by important patient characteristics, such as age, ethnicity, and...ight not generalise to usage in more rural or less affluent areas.
deprivation,
# PRISMA 2020 Checklist

| Section and Topic | Item # | Checklist item | Location where item is reported |
|-------------------|--------|----------------|----------------------------------|
| **TITLE**         | 1      | Identify the report as a systematic review. | 1 ('targeted' as we have adapted our approach) |
| **ABSTRACT**      | 2      | See the PRISMA 2020 for Abstracts checklist. | As per BMJ Open |
| **INTRODUCTION**  | 3      | Describe the rationale for the review in the context of existing knowledge. | 4 |
| **Objectives**    | 4      | Provide an explicit statement of the objective(s) or question(s) the review addresses. | 4 |
| **METHODS**       | 5      | Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses. | 5 |
| **Eligibility criteria** | 6 | Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted. | 4-5 |
| **Information sources** | 7 | Present the full search strategies for all databases, registers and websites, including any filters and limits used. | 4-5 |
| **Search strategy** | 8 | Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process. | 5 |
| **Selection process** | 9 | Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process. | 4-5 |
| **Data collection process** | 10a | List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect. | 5 |
| **Data items** | 10b | List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information. | 5 |
| **Study risk of bias assessment** | 11 | Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process. | n/a |
| **Effect measures** | 12 | Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results. | 5-6 |
| **Synthesis methods** | 13a | Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)). | 5-6 |
| **Effect measures** | 13b | Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions. | 5-6 |
| **Effect measures** | 13c | Describe any methods used to tabulate or visually display results of individual studies and syntheses. | 5-6 |
| **Effect measures** | 13d | Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used. | 5-6 |
| **Effect measures** | 13e | Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression). | 5-6 |
| **Effect measures** | 13f | Describe any sensitivity analyses conducted to assess robustness of the synthesized results. | n/a |
## PRISMA 2020 Checklist

| Section and Topic | Item # | Checklist item | Location where item is reported |
|-------------------|--------|----------------|---------------------------------|
| Reporting bias assessment | 14 | Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases). | n/a |
| Certainty assessment | 15 | Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome. | n/a |

### RESULTS

| Study selection | 16a | Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram. | 6 |
| Study characteristics | 17 | Cite each included study and present its characteristics. | Online suppl. |
| Risk of bias in studies | 18 | Present assessments of risk of bias for each included study. | n/a |
| Results of individual studies | 19 | For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots. | n/a |
| Results of syntheses | 20a | For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies. | Online suppl. |
| | 20b | Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect. | 5-6 |
| | 20c | Present results of all investigations of possible causes of heterogeneity among study results. | 5-6 |
| | 20d | Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results. | 5-6 |
| Reporting biases | 21 | Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed. | n/a |

### DISCUSSION

| Discussion | 23a | Provide a general interpretation of the results in the context of other evidence. | 7-8 |
| | 23b | Discuss any limitations of the evidence included in the review. | 7-8 |
| | 23c | Discuss any limitations of the review processes used. | 7-8 |
| | 23d | Discuss implications of the results for practice, policy, and future research. | 7-9 |

### OTHER INFORMATION

| Registration and protocol | 24a | Provide registration information for the review, including register name and registration number, or state that the review was not registered. | Not registered |
| | 24b | Indicate where the review protocol can be accessed, or state that a protocol was not prepared. | Agreed between authors but not made available publicly. |
### PRISMA 2020 Checklist

| Section and Topic | Item # | Checklist item                                                                                                                                                                                                                                                                                                                                 | Location where item is reported |
|-------------------|--------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------|
| 24c               |         | Describe and explain any amendments to information provided at registration or in the protocol.                                                                                                                                                                                                                                             | n/a                             |
| Support           | 25      | Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.                                                                                                                                                                                                                 | 9                               |
| Competing interests| 26     | Declare any competing interests of review authors.                                                                                                                                                                                                                                                                                          | 9                               |
| Availability of data, code and other materials | 27 | Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.                                                                 | Online Suppl.                   |

*From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71*

*For more information, visit: [http://www.prisma-statement.org/](http://www.prisma-statement.org/)*
## Reporting of data on participant ethnicity and socioeconomic status in high-impact medical journals: a targeted literature review

| Journal: | BMJ Open |
|----------|----------|
| Manuscript ID | bmjopen-2022-064276.R1 |
| Article Type: | Original research |
| Date Submitted by the Author: | 15-Jul-2022 |
| Complete List of Authors: | Buttery, Sara; Imperial College London; NIHR Imperial Biomedical Research Centre Philip, Keir; Imperial College London; NIHR Imperial Biomedical Research Centre Alghamdi, Saeed; Imperial College London Williams, Parris; Imperial College London, National Heart and Lung Institute Quint, Jennifer; Imperial College London, NHLI; Imperial College London Hopkinson, Nicholas; Imperial College London, National Heart and Lung Institute |
| **Primary Subject Heading**: | Medical publishing and peer review |
| **Secondary Subject Heading**: | Medical publishing and peer review |
| **Keywords**: | STATISTICS & RESEARCH METHODS, GENERAL MEDICINE (see Internal Medicine), INTERNAL MEDICINE |
I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd (“BMJ”) its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge (“APC”) for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author’s Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.
Reporting of data on participant ethnicity and socioeconomic status in high-impact medical journals: a targeted literature review

Sara C Buttery\textsuperscript{1,2} BSc*, Keir EJ Philip\textsuperscript{1,2} MRCP*, Saeed M. Alghamdi\textsuperscript{1,2} MSc, Parris Williams\textsuperscript{1,2} MSc, Jennifer K Quint\textsuperscript{1,2} PhD**, Nicholas S Hopkinson\textsuperscript{1,2} PhD**

*Joint first authors

**Joint senior authors

Affiliations:

1) National Heart and Lung Institute, Imperial College London, London, United Kingdom

2) NIHR Imperial Biomedical Research Centre, London, United Kingdom

Correspondence to:

Dr Keir EJ Philip, The Muscle Laboratory, The National Heart and Lung Institute, Imperial College London, London, NW3 6HP. \texttt{K.Philip@imperial.ac.uk}

Keywords: ethnicity, race, socioeconomic status, participant characteristics, demographics

Word count: 1716
ABSTRACT

Objectives: To assess the frequency of reporting of ethnicity (or 'race') and socioeconomic status (SES) indicators in high-impact journals.

Design: Targeted literature review.

Data sources: The 10 highest ranked general medical journals using Google scholar h5 index.

Eligibility criteria: Inclusion criteria were, human research, reporting participant level data.

Exclusion criteria were non-research article, animal/other non-human participant/subject; or no participant characteristics reported.

Data extraction and synthesis: Working backwards from April 19th, 2021 in each journal, two independent reviewers selected the 10 most recent articles meeting inclusion/exclusion criteria, to create a sample of 100 articles. Data on the frequency of reporting of ethnicity (or ‘race’) and SES indicators were extracted and presented using descriptive statistics.

Results: Of one hundred research articles included, 35 reported ethnicity and 13 SES. By contrast, 99 reported age, and 97 reported sex or gender. Among the articles not reporting ethnicity only 3 (5%) highlighted this as a limitation, and only 6 (7%) where SES data were missing. Median number of articles reporting ethnicity per journal was 2.5/10 (range 0 to 9). Only 2 journals explicitly requested reporting of ethnicity (or race), and 1 requested SES.

Conclusions: The majority of research published in high-impact medical journals does not include data on the ethnicity and socioeconomic status of participants, and this omission is rarely acknowledged as a limitation. This situation persists despite the well-established importance of this issue and ICMJE recommendations to include relevant demographic variables to ensure representative samples. Standardized explicit minimum standards are required.
**Strengths and limitations of this study**

- This study included recent studies from a range of the highest impact general medical journals.

- Different inclusion/exclusion criteria for articles could be justifiably used, which may have produced different results.

- We identified high-impact journals using the google scholar h5 index, however various other equally valid impact metrics exist, which could change the journals considered.

- Our analysis focused on *if* ethnicity and/or race was reported, but not *how* they are reported which is an important and related area for discussion and research to that covered in this study.
Introduction

Information about the ethnicity and socioeconomic status of participants in clinical research is needed for the interpretation, generalisability and pooling of data as well as to inform discussion around health inequalities. The relevance of ethnicity and socio-economic status to health and biomedical research is well established but has been emphasised by the COVID-19 pandemic, during which specific ethnic groups and poorer individuals have been disproportionately affected. The causal pathways driving health disparities are complex and multifactorial, however under-reporting of participant characteristics has been identified as a potential contributory factor.

The International Committee of Medical Journal Editors recommendations, and some journal instructions to authors promote inclusion of these data. Previous studies have identified that reporting is frequently incomplete with limited progress made over the last three decades. Recent years have seen an increased focus on ethnicity and socioeconomic status in medicine, however there is a lack of research as to whether this has resulted in better reporting.

To evaluate the current situation in this area, we assessed the frequency of reporting of ethnicity (or ‘race’) and socioeconomic status indicators in a sample of research articles published in high impact general medical journals in Spring 2021.

Methods

We identified the 10 highest ranked journals as per Google scholar ‘Health and Medical (general)’ category up to April 2021. At the time of data collection these were The New England Journal of Medicine (NEJM), The Lancet, the Journal of the American Medical Association, Proceedings of the National Academy of Sciences of the United States of America (PNAS), Nature Medicine, Public Library of Science One (PLOS One), The British Medical Journal (BMJ), Cochrane, Cell Metabolism, and Science Translational Medicine. PNAS and PLOS One include a wide range of subject areas therefore the subsections ‘Biological Sciences, Medical Science’ and ‘Clinical Medicine’ were used.
respectively. From each of these 10 journals, using the journals’ own websites, we worked backwards from April 19th, 2021, selecting the 10 most recent journal articles that met inclusion/exclusion criteria. Inclusion criteria were: research articles, reporting participant level data. Articles were excluded if they were: not research (e.g. editorial, news, images etc.), animal/other non-human participant/subject; or no participant characteristics reported. Laboratory studies using human derived tissues or cells were included if donor information was provided. Journal reporting guidance and requirements were also assessed by evaluating author guidelines, websites, and contacting the respective editorial/publishing teams. Data were collected on which participant level characteristics were reported and how. Data were also collected on if the absence of reporting these variables was noted as a limitation. The journals’ accessible policies and guidance on reporting these variables was also reviewed. Data collection and analysis was conducted by SCB, KEJP, SMA and PW. All journals were reviewed and articles selected by at least two researchers independently, who then came together to discuss any inconsistencies with a third researcher.

Ethnicity and race are related yet different constructs and arguably the latter term should be abandoned. However, given the frequent lack of standardisation in the literature and that the terms are in practice often used interchangeably we accepted the use of either term. For the purpose of this study ethnicity (or race) was defined as variables explicitly stated by the authors as ‘ethnicity’, ‘ethnic group’, or ‘race’, ‘racial group’. Similarly, regarding reporting of socioeconomic status indicators, various often inconsistent methods are used, therefore we opted to assess both direct measures such as the Index of Multiple Deprivation, but also measures from which socioeconomic status could be inferred such as educational attainment and job role. The focus being if, rather than how, such measures are reported. Variables were considered to be indicators of SES if they were explicitly stated as being included for this purpose in the studies reporting them, or if not explicitly stated in the study itself, variables that might be considered SES indicators were discussed between researchers and included or excluded based on consensus opinion. Given the potential degree of subjectivity related to this approach we have provided the specific terms used by included studies in the results section below. The agreed approach
was to take a more inclusive approach, so that if these variables were found to be infrequently reported, such findings would not be dismissed as relating to overly stringent inclusion criteria.

Patient and public involvement

None.

Results

650 publications were assessed to identify 100 meeting inclusion criteria (see figure 1 and Supplementary Material Tables 1, 2, and 3). Of one hundred research articles included, 35 reported ethnicity (or race) and 13 reported socioeconomic status. By contrast, 99 reported age, and 97 reported sex or gender (Table1).

Among the articles not reporting ethnicity only 3 (5%) highlighted this as a limitation, and only 6 (7%) highlighted where socioeconomic status data were missing. Median number of articles reporting ethnicity per journal was 2.5/10 (range 0/10 (PLOS One), to 9/10 ). Only 2 journals explicitly requested reporting of participant ethnicity (or race), and 1 requested socioeconomic status. Types of research included – interventional studies (n=30), cohort studies (n=35), case-control studies (n=3), systematic reviews and metanalyses (n=16), epidemiological and surveys (n=3), and other (n=13). Twenty of the 100 were laboratory studies (either observational or involving interventional manipulation of samples) using human samples, of which 4 reported ethnicities of sample donors (of others, none mentioned as a limitation), and none reported socioeconomic status.

Among the 24 papers describing clinical trials, 50% reported ethnicity, with none highlighting the absence of these data as a limitation. 12.5% of trials reported an indicator of socioeconomic status, with one of the 21 not reporting socioeconomic status highlighting this absence as a limitation.

Of note, two of the research articles included in our sample identified ethnicity as being relevant to their research topic, yet did not provide relevant data on their study participants or highlight the lack of this
data as a limitation of their study ‘in the case of DNA-based mutation testing, poor sensitivity in detecting mutations in infants from ethnic and racial minority groups’, and ‘peripheral oxygen saturation can substantially differ from the SaO\textsubscript{2} under certain conditions and may be less accurate in Black patients than in White patients.’\textsuperscript{15}.

**Figure 1: Flow diagram of included/excluded articles**
Table 1: Reporting of ethnicity and/or race, and Socioeconomic Status indicators in research articles

| Report participant level characteristics | N | Additional notes |
|-------------------------------------------|---|-----------------|
| Report ethnicity and/or race              |   |                 |
| 35/100 report                             | 65 Not report | Range per journal: JAMA 9/10, with clear guidance that this information is expected. |
| Noted in limitations                      |   | Some studies identify race and ethnicity as being relevant to the research focus, yet did not provide relevant data on their study participants or highlight this a limitation of their study e.g. |
| 62 of the 65 do not state this as a limitation | 3 Do highlight this as a limitation. | - 'in the case of DNA-based mutation testing, poor sensitivity in detecting mutations in infants from ethnic and racial minority groups' (DOI: 10.1126/scitranslmed.abd8109) |
|                                          |   | - 'peripheral oxygen saturation can substantially differ from the Sao2 under certain conditions and may be less accurate in Black patients than in White patients.' (DOI: 10.1056/NEJMoa2032510) |
| Report socioeconomic status indicator     |   |                 |
| 13/100 report at a measure of SES (6 direct measure e.g. Index of Multiple Deprivation, Poverty income ratio; 7 measures from which SES can be inferred e.g educational attainment, job role) | 87/100 did not report any indication of SES |
| Noted in limitations                      |   |                 |
| 6/87 identified this as a limitation      |   |                 |
| Age reported                              | 99/100 |   |
| Sex or Gender reported                   | 97/100 |   |

Percentages not given as most results have 100 as the denominator.
Discussion

The majority of research published in high-impact medical journals does not include data on the ethnicity and socioeconomic status of participants, and this omission is rarely acknowledged as a limitation. This finding echoes related historical research, but its persistence is of concern and is surprising given current awareness of such issues.

These findings have important implications for the interpretation and application of research findings, both within academia and beyond, with the ongoing omission no longer justifiable as simple oversight. As highlighted by Baker et al. in relation to data relating to LGBTQI+ communities, but equally relevant here, ‘Data are fundamentally political: decisions about which data are collected and which are overlooked both reflect and shape policy and program priorities.’

Our results could have multiple contributory factors. For some research including secondary data analyses, ethnicity and socioeconomic status data may not have been available to the researchers, but given the lack of explanation, it remains unclear if these data were unavailable, or available but not included in publications. The low level of reporting in controlled clinical trials suggests issues beyond unavailability of data, as in these studies such data would be simple to collect. Additionally, given research successfully reporting these data, the justification for these omissions remains unexplained. Non-reporting of ethnicity (or race) and SES data may also result from explicit or implicit racism, or other forms of discrimination such as that based on SES, which could include failing to appreciate the relevance of these factors to the generalisability of findings.

The increased frequency of reporting ethnicity compared to socioeconomic status, may indicate differences between the perceived relevance of these variables. This would be in keeping journal author guidelines and ICMJE recommendations that encourage the inclusion of relevant demographic variables to ensure representative samples, more often explicitly stating race and/or ethnicity, than socioeconomic status. The relevance of these factors may not have been apparent to authors and editorial teams, however ICMJE Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly work in Medical Journals states ‘Because the relevance of such variables as
age, sex, or ethnicity is not always known at the time of study design, researchers should aim for inclusion of representative populations into all study types and at a minimum provide descriptive data for these and other relevant demographic variables.’. Of note, not all of the journals in our sample state that they follow the ICMJE recommendations. However, whether or not the journal states they follow guidance or not, this has no impact upon the relevance of these data and the importance of reporting them. Additionally, Maduka et al found no difference between journals stating they follow ICMJE recommendations, and those that do not, in the frequency of reporting race and ethnicity in a sample of surgical research publications in 2019.

Certain considerations and limitations require highlighting. Firstly, different approaches to selecting research papers may alter findings. Secondly, we identified high-impact journals using the google scholar h5 index but acknowledge various other equally valid methods exist. Thirdly our analysis focused on if ethnicity and/or race was reported, but we acknowledge that these are not synonymous terms. In addition to if these variables are reported, how they are reported is also an important area for discussion and research. The choice to analyse 100 papers was somewhat arbitrary. We wanted to include an adequate number of articles from the selected journals to provide a representative sample of their original research papers. Furthermore, given the substantial differences in the number of original research papers published between journals, keeping to ten per journal ensured all included papers were published within a 4-month window. If we had included 100 papers per journal, the sample from some journals might be 2 months, while others nearer 2 years, which could complicate interpretation given the potential for changing levels of reporting over time. The widespread omissions identified by this research suggests a structural problem. Indeed, we the authors have published research which would have met the inclusion criteria and failed to report these specific characteristics. Our intention is to highlight an issue and suggest approaches to address it.

Given that inadequate reporting persists despite research highlighting the issue, author and ICMJE recommendations, and the current socio-political climate, there is a clear need for more explicit requirements that are adhered to in practice. This is likely best achieved if steps are integrated into each stage of the research process, from protocol to publication. For example, Fain et al compared reporting
of race and ethnicity on ClinicalTrials.gov before and after the requirement to report these data (if collected), was introduced, finding that this was associated with an increase from 42% to 92%. Similar explicit requirements could be taken in EQUATOR guidelines\textsuperscript{22}, and research ethics applications. From our sample, the journal JAMA had the most explicit guidance for reporting race and ethnicity, and this variable was reported in 9/10 of the articles we reviewed. Of note, from 2022 the New England Journal of Medicine will be requiring authors of research articles to provide data on the representativeness of the sample including race or ethnic group\textsuperscript{23}, though it is unclear if socioeconomic status indicators will also be required. Much of the recent literature appears to focus on ethnicity reporting, likely due to the COVID-19 pandemic exposing its disproportionate effects on some ethnic groups \textsuperscript{24}. One recent publication in Nature medicine\textsuperscript{24} suggested it would require changes at policy level as well as engaging with professionals, patients and the public to communicate the importance of this issue in understanding inequalities, Barriers suggested include problems collecting ethnicity data, whether this be reported by a healthcare professional or self-reported, and in defining ethnic groups where categorisation is inconsistent. \textsuperscript{24}\textsuperscript{25} This is reflected in the diverse terms used to report ethnicity in the papers we reviewed (Table 3 Supplementary Material). Future research would be useful investigating changing in reporting overtime, especially in relation to specific actions taken to improve this issue, which could inform research reporting guidelines.

**Conclusion**

The reporting of ethnicity and socioeconomic status in high-impact medical research remains poor, despite a consensus on its importance. Omission of these participant characteristics limits the interpretation, generalisability, and pooling of data, that are required to facilitated informed discussion around health inequalities. Guidance and encouragement have so far proven insufficient to change practice in this area. Standardised, explicit, minimum standards are required.
Contributors

SCB, had the original idea for the study. SCB, KEJP, SMA and PW collected the data. All authors (SCB, KEJP, SMA, PW, JKQ and NSH) contributed to the design of the study. KEJP analysed the data initially, which was verified by SCB, SMA and PW. KEJP wrote the first draft of the manuscript. All authors (SCB, KEJP, SMA, PW, JKQ and NSH) critically appraised the manuscript and approved it for submission and had full access to the data and can take responsibility for the integrity of the data and the accuracy of the data analysis. The corresponding author attests that all listed authors (SCB, KEJP, SMA, PW, JKQ and NSH) meet authorship criteria and that no others meeting the criteria have been omitted.

Funding

KEJP is supported was supported by the Imperial College Clinician Investigator Scholarship (internal award with no specific grant number/code). The funders had no say in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Competing interests

None reported.

Acknowledgments

None.

Data availability statement

All data used in this study are publicly available.

Ethics approval

Ethics approval for this study was not required as all data used are freely available in the published literature.
References

1. Williamson EJ, Walker AJ, Bhaskaran K, et al. Factors associated with COVID-19-related death using OpenSAFELY. *Nature* 2020;584(7821):430-36. doi: 10.1038/s41586-020-2521-4 [published Online First: 2020/07/09]

2. Chastain DB, Osae SP, Henao-Martinez AF, et al. Racial Disproportionality in Covid Clinical Trials. *The New England journal of medicine* 2020;383(9):e59. doi: 10.1056/NEJMp2021971 [published Online First: 2020/08/12]

3. Krieger N, Waterman PD, Chen JT, et al. Missing again: US racial and ethnic data for COVID-19 vaccination. *Lancet* 2021;397(10281):1259-60. doi: 10.1016/S0140-6736(21)00465-7 [published Online First: 2021/03/21]

4. Webb Hooper M, Napoles AM, Perez-Stable EJ. COVID-19 and Racial/Ethnic Disparities. *Jama* 2020;323(24):2466-7. doi: 10.1001/jama.2020.8598 [published Online First: 2020/05/12]

5. ICMJE. Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly work in Medical Journals. Updated December 2019, 2019.

6. The Lancet. Information for Authors 2021 [cited 2021 02/09]. Available from: https://marlin-prod.literatumonline.com/pb-assets/Lancet/authors/tl-info-for-authors.pdf.

7. JAMA. Instructions for Authors 2021 [cited 2021 02/09]. Available from: https://jamanetwork.com/journals/jama/pages/instructions-for-authors.

8. Geller SE, Koch A, Pellettieri B, et al. Inclusion, analysis, and reporting of sex and race/ethnicity in clinical trials: have we made progress? *J Womens Health (Larchmt)* 2011;20(3):315-20. doi: 10.1089/jwh.2010.2469 [published Online First: 2011/03/01]

9. Brahan D, Bauchner H. Changes in reporting of race/ethnicity, socioeconomic status, gender, and age over 10 years. *Pediatrics* 2005;115(2):e163-6. doi: 10.1542/peds.2004-1437 [published Online First: 2005/02/03]

10. Kanakamedala P, Haga SB. Characterization of clinical study populations by race and ethnicity in biomedical literature. *Ethn Dis* 2012;22(1):96-101. [published Online First: 2012/07/11]

11. Bokor-Billmann T, Langan EA, Billmann F. The reporting of race and/or ethnicity in the medical literature: a retrospective bibliometric analysis confirmed room for improvement. *Journal of clinical epidemiology* 2020;119:1-6. doi: 10.1016/j.jclinepi.2019.11.005 [published Online First: 2019/11/13]

12. Lee SJ, Kavanaugh A. A need for greater reporting of socioeconomic status and race in clinical trials. *Ann Rheum Dis* 2004;63(12):1700-1. doi: 10.1136/ard.2003.019588 [published Online First: 2004/11/18]

13. Ma IW, Khan NA, Kang A, et al. Systematic review identified suboptimal reporting and use of race/ethnicity in general medical journals. *Journal of clinical epidemiology* 2007;60(6):572-8. doi: 10.1016/j.jclinepi.2006.11.009 [published Online First: 2007/05/12]

14. Ioannidis JPA, Powe NR, Yancy C. Recalibrating the Use of Race in Medical Research. *Jama* 2021;325(7):623-24. doi: 10.1001/jama.2021.0003 [published Online First: 2021/01/26]

15. Schjorring OL, Klitgaard TL, Perner A, et al. Lower or Higher Oxygenation Targets for Acute Hypoxic Respiratory Failure. *The New England journal of medicine* 2021;384(14):1301-11. doi: 10.1056/NEJMoa2032510 [published Online First: 2021/01/21]

16. Flanagin A, Frey T, Christiansen SL, et al. The Reporting of Race and Ethnicity in Medical and Science Journals: Comments Invited. *Jama* 2021;325(11):1049-52. doi: 10.1001/jama.2021.2104 [published Online First: 2021/02/23]

17. The New England Journal of Medicine. Race and Medicine 2021 [cited 2021 05/09/2021].

18. Baker KE, Streed CG, Jr., Durso LE. Ensuring That LGBTQ+ People Count - Collecting Data on Sexual Orientation, Gender Identity, and Intersex Status. *The New England journal of medicine* 2021;384(13):1184-86. doi: 10.1056/NEJMep2032447 [published Online First: 2021/04/02]
19. International Committee of Medical Journal Editors. Journals stating that they follow the ICMJE Recommendations 2021 [cited 2021 09/09/2021]. Available from: http://www.icmje.org/journals-following-the-icmje-recommendations/.

20. Maduka RC, Broderick M, White EM, et al. The Reporting of Race and Ethnicity in Surgery Literature. *JAMA Surg* 2021 doi: 10.1001/jamasurg.2021.3752 [published Online First: 2021/08/19]

21. Fain KM, Nelson JT, Tse T, et al. Race and ethnicity reporting for clinical trials in ClinicalTrials.gov and publications. *Contemp Clin Trials* 2021;101:106237. doi: 10.1016/j.cct.2020.106237 [published Online First: 2020/12/09]

22. EQUATOR network. Reporting guidelines 2021 [Available from: https://www.equator-network.org/reporting-guidelines/ accessed 04/10/2021 2021.

23. Rubin E. Striving for Diversity in Research Studies. *The New England journal of medicine* 2021 doi: 10.1056/NEJMe2114651 [published Online First: 2021/09/14]

24. Routen A, Akbari A, Banerjee A, et al. Strategies to record and use ethnicity information in routine health data. *Nature Medicine* 2022 doi: 10.1038/s41591-022-01842-y

25. Khunti K, Routen A, Banerjee A, et al. The need for improved collection and coding of ethnicity in health research. *J Public Health (Oxf)* 2021;43(2):e270-e72. doi: 10.1093/pubmed/fdaa198

**Figure titles**

**Figure 1:** Flow diagram of study inclusion/exclusion
Figure 1: Flow diagram of study inclusion/exclusion

Journals with the highest Google Scholar h5 index: NEJM, The Lancet, JAMA, PNAS, Nature Medicine, PLOS One, BMJ, Cochrane, Cell Metabolism, and Science Translational Medicine

Working backwards from April 19th 2021 through each journal until 10 research articles included meeting inclusion criteria. 650 publications assessed in total to reach target number of articles.

Excluded (n= 550)
- Non-research (n=442)
- Animal/other non-human participant/subject (n=83)
- No participant characteristics reported (n=25)

100 research publications included (x10 per journal)
- 88 Full articles
- 12 Research Letters/Brief Communications (All additional supplementary data also reviewed)
Table 1: Research papers included in the sample

| Journal | Date of pub | Title                                                                 | DOI                          | Country of journal      | Country of study/Corresponding author | Manuscript type | Study design | Report baseline/participants characteristics (which & how) |
|---------|-------------|-----------------------------------------------------------------------|------------------------------|--------------------------|---------------------------------------|-----------------|--------------|-----------------------------------------------------------|
| NEJM    | 15/04/2021  | Hypothermic Machine Perfusion in Liver Transplantation — A Randomized Trial | 10.1056/NEJMoa2031532        | USA                      | Multicentre Europe                     | Original research (full paper) | RCT          | Yes: Age, male sex, BMI, preservation of liver measures,   |
| NEJM    | 15/04/2021  | Trial of Psilocybin versus Escitalopram for Depression                  | 10.1056/NEJMoa2032994        | USA                      | UK                                    | Original research (full paper) | RCT          | Yes: Age, female sex, white race, employment status,       |
| NEJM    | 15/04/2021  | BNT162b2 mRNA Covid-19 Vaccine in a Nationwide Mass Vaccination Setting | 10.1056/NEJMoa2101765        | USA                      | Israel                                | Original research (full paper) | Case Control | Yes: Age, female/male sex, population sector (general Jewish, Arab, Ultra-orthodox Jewish), comorbidities, |
| NEJM    | 15/04/2021  | Dexmedetomidine or Propofol for Sedation in Mechanically Ventilated Adults with Sepsis | 10.1056/NEJMoa2024922        | USA                      | USA                                   | Original research (full paper) | RCT          | Yes: Age, Female sex %, BMI, 'Race or Ethnic Group' White, Black, Latinx, multiple or other; cognitive decline score; clinical illness |
| NEJM    | 08/04/2021  | Lenvatinib plus Pembrolizumab or Everolimus for Advanced Renal Cell Carcinoma | 10.1056/NEJMoa2035716        | USA                      | Global                                | Original research (full paper) | RCT          | Yes: Age, sex (male/female), geographic region          |
| NEJM    | 08/04/2021  | Lower or Higher Oxygenation Targets for Acute Hypoxemic Respiratory Failure | 10.1056/NEJMoa2032510        | USA                      | Denmark                               | Original research (full paper) | RCT          | Yes: Age, sex %male, comorbidities, illness/admission metrics |
| NEJM    | 08/04/2021  | Glycemic Index, Glycemic Load, and Cardiovascular Disease and Mortality | 10.1056/NEJMoa2007123        | USA                      | Global                                | Original research (full paper) | Cohort study | Yes: Age, sex %male, urban residence, health risk factors, results by continents |
| NEJM    | 08/04/2021  | Sutimlimab in Cold Agglutinin Disease                                  | 10.1056/NEJMoa2027760        | USA                      | Germany                               | Original research (full paper) | Intervention trial (other than RCT) | Yes: Age, sex %female, geographic location (Europe, Japan, USA, Australia), disease characteristics, |
| Journal | Date       | Title                                                                 | DOI                          | Country     | Country | Study Type            | Details                                                                 |
|---------|------------|----------------------------------------------------------------------|------------------------------|-------------|----------|-----------------------|------------------------------------------------------------------------|
| NEJM    | 08/04/2021 | Antibody Responses in Seropositive Persons after a Single Dose of SARS-CoV-2 mRNA Vaccine | 10.1056/NEJMc2101667         | USA         | USA      | Original research (letter) | Cohort study; Yes: Age, gender (male, female), prefer not to say,          |
| NEJM    | 01/04/2021 | Adjuvant Nivolumab in Resected Esophageal or Gastric Esophageal Junction Cancer | 10.1056/NEJMoa2032125        | USA         | Global   | Original research (full paper) | RCT; Yes: Age, male sex %, race (white, Asian, black, other, not reported), Geographic region (Europe, US, Canada, Asia) |
| The Lancet | 17/04/2021 | Thromboembolism and the Oxford–AstraZeneca COVID-19 vaccine: side-effect or coincidence? | 10.1016/S0140-6736(21)00762-5 | UK          | Denmark  | Original research (letter) | Cohort study; Yes: age group, female + male numbers                      |
| The Lancet | 17/04/2021 | Effect of infusion set replacement intervals on catheter-related bloodstream infections (RSVP): a randomised, controlled, equivalence (central venous access device)–non-inferiority (peripheral arterial catheter) trial | 10.1016/S0140-6736(21)00351-2 | UK          | Australia| Original research (full paper) | RCT; Yes: Gender (female, male, other); Age; Ethnicity (white, mixed race, Asian, black, Chinese, other, prefer not to say), medical conditions, index of multiple deprivation, region of England. |
| The Lancet | 17/04/2021 | SARS-CoV-2 infection rates of antibody-positive compared with antibody-negative health-care workers in England: a large, multicentre, prospective cohort study (SIREN) | 10.1016/S0140-6736(21)00675-9 | UK          | England  | Original research (full paper) | Cohort study                                                              |
| The Lancet | 10/04/2021 | Efficacy of ChAdOx1 nCoV-19 (AZD1222) vaccine against SARS-CoV-2 variant of concern 202012/01 (B.1.1.7): an exploratory analysis of a randomised controlled trial | 10.1016/S0140-6736(21)00628-0 | UK          | UK       | Original research (full paper) | RCT (Secondary data analysis); Yes: age, % female, ethnicity white, black, Asian, mixed, other, missing, |
| The Lancet | 10/04/2021 | The SANAD II study of the effectiveness and cost-effectiveness of levetiracetam, zonisamide, or lamotrigine for newly diagnosed focal epilepsy: an open-label, non-inferiority, multicentre, phase 4, randomised controlled trial | 10.1016/S0140-6736(21)00247-6 | UK          | UK       | Original research (full paper) | RCT; Yes: Age, gender (male/female),                                       |
| The Lancet | 10/04/2021 | The SANAD II study of the effectiveness and cost-effectiveness of valproate versus levetiracetam for newly diagnosed generalised and unclassifiable epilepsy: an open-label, non- | 10.1016/S0140-6736(21)00246-4 | UK          | UK       | Original research (full paper) | RCT; Yes: Age, gender (male/female),                                       |
| Journal       | Date       | Study Title                                                                 | DOI                                                                 | Country         | Disease variables                                         | Study Design | Description                                                                 |
|--------------|------------|------------------------------------------------------------------------------|---------------------------------------------------------------------|-----------------|------------------------------------------------------------|--------------|-----------------------------------------------------------------------------|
| The Lancet   | 03/04/2021 | Efficacy and safety of dolutegravir with emtricitabine and tenofovir alafenamide fumarate or tenofovir disoproxi fumarate, and efavirenz, emtricitabine, and tenofovir disoproxi fumarate HIV antiretroviral therapy regimens started in pregnancy (IMPAACT 2010/VESTED): a multicentre, open-label, randomised, controlled, phase 3 trial | 10.1016/S0140-6736(21)003314-7                                     | UK              | Original research (full paper)                             | RCT          | Yes: Age, all female (in pregnancy), Country, race (Black, Asian, White, Other, unknown), |
| The Lancet   | 03/04/2021 | Comparison of two delayed strategies for renal replacement therapy initiation for severe acute kidney injury (AKIKI 2): a multicentre, open-label, randomised, controlled trial | 10.1016/S0140-6736(21)00350-0                                    | France          | Original research (full paper)                             | RCT          | Yes: Age, sex (female/male), comorbidities                                 |
| The Lancet   | 27/03/2021 | Evaluating Progestogens for Preventing Preterm birth International Collaborative (EPPPIC): meta-analysis of individual participant data from randomised controlled trials | 10.1016/S0140-6736(21)00217-8                                    | Not provided    | Original research (full paper)                             | Systematic review and meta-analysis | Yes: age, all female (in pregnancy) ethnicity (Black, Asian, Hispanic, middle eastern, other, white, unknown), disease variables |
| The Lancet   | 27/03/2021 | Discontinuing β-lactam treatment after 3 days for patients with community-acquired pneumonia in non-critical care wards (PTC): a double-blind, randomised, placebo-controlled, non-inferiority trial | 10.1016/S0140-6736(21)003313-5                                    | France          | Original research (full paper)                             | RCT          | Yes: Age, sex (female/male), age, race (white, black or African American, other, Asian, Native Hawaiian or other pacific island, American Indian or Alaska native, Hispanic or Latino ethnic group, body weight, BMI, comorbidities, clinical measurements, |
| JAMA         | 13/04/2021 | Effect of Subcutaneous Semaglutide vs Placebo as an Adjunct to Intensive Behavioural Therapy on Body Weight in Adults With Overweight or Obesity The STEP 3 Randomized Clinical Trial | 10.1001/jama.2021.1831                                            | USA             | Original research (full paper)                             | RCT          | Yes: age, sex (women, men), race (white, black or African American, other, Asian, Hispanic or Latino ethnic group, body weight, BMI), comorbidities, clinical measurements, |
| JAMA         | 13/04/2021 | Effect of Continued Weekly Subcutaneous Semaglutide vs Placebo on Weight Loss Maintenance in Adults With | 10.1001/jama.2021.3224                                            | Global          | Original research (full paper)                             | RCT          | Yes: age, sex (women, men), race (white, black or African American, other, Asian, Hispanic or Latino ethnic group, body weight, BMI), comorbidities, clinical measurements, |
| Journal | Date | Title | Study Details | Type | Design | Key Variables |
|---------|------|-------|---------------|------|--------|---------------|
| JAMA    | 13/04/2021 | Overweight or Obesity The STEP 4 Randomized Clinical Trial | comorbidities, clinical measurements, | Original research (full paper) | RCT | Yes: age, sex (male, female), race or ethnic group (mixed race, Black or African American, Colombian native), Health Insurance (private/semiprivate, government subsidised, uninsured), number of people in the household, current smoker, BMI, Comorbidities etc |
| JAMA    | 13/04/2021 | Effect of Ivermectin on Time to Resolution of Symptoms Among Adults with Mild COVID-19A Randomized Clinical Trial | USA Colombia | Original research (full paper) | RCT | Yes: age, sex (male, female), race/ethnicity (Black or Black American, White, Asian) vaccine received |
| JAMA    | 13/04/2021 | Binding and Neutralization Antibody Titers After a Single Vaccine Dose in Health Care Workers Previously Infected With SARS-CoV-2 | USA USA | Original research (Letter) | Cohort study | Yes: age, sex (male, female), race/ethnicity (Non-Hispanic white, Hispanic, Native American, Black), BMI, comorbidities, lab results |
| JAMA    | 06/04/2021 | Effect of Low-Intensity vs High-Intensity Home-Based Walking Exercise on Walk Distance in Patients With Peripheral Artery Disease The LITE Randomized Clinical Trial | USA USA | Original research (full paper) | RCT | Yes: Age, Sex (Male/Female), Race White, Black, Asian, Other, Hispanic ethnicity. |
| JAMA    | 06/04/2021 | Effect of Celecoxib vs Placebo Added to Standard Adjuvant Therapy on Disease-Free Survival Among Patients With Stage III Colon Cancer The CALGB/SWOG 80702 (Alliance) Randomized Clinical Trial | USA USA | Original research (full paper) | RCT | Yes: Age, Sex (Men/Women), Race (White, Black or African American, Asian, All others or not reported), Hispanic or Latino %) Disease characteristics |
| JAMA    | 06/04/2021 | Antimicrobial Use in a Cohort of US Nursing Homes, 2017 | USA USA | Original research (full paper) | Cohort study | Yes: Age, sex (men/women), race/ethnicity (Other, Hispanic or Latino, Black non-Hispanic, white non-Hispanic,) |
| JAMA    | 06/04/2021 | Trends in Age at Natural Menopause and Reproductive Life Span Among US Women, 1950-2018 | USA USA | Original research (Letter) | Epidemiologic assessment survey | Yes: Age, (all female), Race/ethnicity (White, Black, Hispanic, non-US born), Educational attainment,
| Journal | Date      | Title                                                                 | DOI                        | Country(s) | Study Type                  | Other Comments                                                                 |
|---------|-----------|----------------------------------------------------------------------|----------------------------|------------|-----------------------------|--------------------------------------------------------------------------------|
| JAMA    | 30/03/21  | Intubation Practices and Adverse Peri-intubation Events in Critically Ill Patients From 29 Countries | 10.1001/jama.2021.1727     | USA, Global | Original research (full paper) | Other observational study, Yes: Age, Women%, comorbidities |
| PNAS    | 30/03/21  | Estrogen receptor β and treatment with a phytoestrogen are associated with inhibition of nuclear translocation of EGFR in the prostate | 10.1073/pnas.2011269118     | USA, Sweden | Original research (full paper) | Cohort study (lab), Yes, Sex (all Males), age, ethnicity |
| PNAS    | 30/03/21  | Health and economic impact of the pneumococcal conjugate vaccine in hindering antimicrobial resistance in China | 10.1073/pnas.2004933118      | USA, China  | Original research (full paper) | Other (Mathematical Modelling) Yes age |
| PNAS    | 16/03/21  | Loss of expression of both miR-15/16 loci in CML transition to blast crisis | 10.1073/pnas.2101566118      | USA, USA    | Original research (full paper) | Cohort study (lab), Yes, sex |
| PNAS    | 09/03/21  | Influence of a COVID-19 vaccine’s effectiveness and safety profile on vaccination acceptance | 10.1073/pnas.2021726118      | USA, USA    | Original research (full paper) | Survey Yes, sex, age, race |
| PNAS    | 09/03/21  | Elevated cerebrospinal fluid cytokine levels in tuberculous meningitis predict survival in response to dexamethasone | 10.1073/pnas.2024852118      | USA, USA    | Original research (full paper) | Cohort study (lab/modelling) Yes age, (sex/gender not reported) |
| PNAS    | 02/03/21  | Glucagon blockade restores functional β-cell mass in type 1 diabetic mice and enhances function of human islets | 10.1073/pnas.2022142118      | USA, USA    | Original research (full paper) | Interventional (lab) Yes Sex, Age |
| PNAS    | 23/03/21  | Modelling SARS-CoV-2 viral kinetics and association with mortality in hospitalized patients from the French COVID cohort | 10.1073/pnas.2017962118      | USA, France | Original research (full paper) | Cohort study Yes, Gender, Age |
| PNAS    | 09/02/21  | Arsenic trioxide replacing or reducing chemotherapy in consolidation therapy for acute promyelocytic leukemia (APL2012 trial) | 10.1073/pnas.2020382118      | USA, China  | Original research (full paper) | RCT Yes, age, sex |
| PNAS    | 02/02/21  | Efficient detection and post-surgical monitoring of colon cancer with a multi-marker DNA methylation liquid biopsy | 10.1073/pnas.2017421118      | USA, China  | Original research (full paper) | Cohort study Yes, Sex, age |
| PNAS    | 05/01/21  | A data-driven approach to identify risk profiles and protective drugs in COVID-19 | 10.1073/pnas.2016877118      | USA, Switzerland | Original research (full paper) | Cohort study Yes, age, sex |
| Journal                  | Date         | Title                                                                 | DOI                          | Country/Country | Study Type                  | Study Design                     | Notes                                                                 |
|-------------------------|--------------|-----------------------------------------------------------------------|------------------------------|-----------------|-----------------------------|----------------------------------|----------------------------------------------------------------------|
| Nature Medicine         | 15/04/2021   | Integrative microbiomics in bronchiectasis exacerbations              | 10.1016/S0140-6736(21)00313-5 | US/Scotland     | Original research (full paper) | Cohort study                     | yes; age, gender, geographic origin, aetiology, smoking status, BSI (status/score), BMI, MRC, FEV1 |
| Nature Medicine         | 15/04/2021   | Assessment of medication self-administration using artificial intelligence | 10.1038/s41591-021-01273-1    | US/Kosovo       | Original research (full paper) | Cohort study                     | yes; gender, and Age                                                      |
| Nature Medicine         | 15/04/2021   | Malaria is a cause of iron deficiency in African children             | 10.1038/s41591-021-01238-4    | Africa          | Original research (brief communication/letter) | Cohort study                     | yes; age, gender (female), inflammation, underweight                  |
| Nature Medicine         | 15/04/2021   | Attributes and predictors of long COVID                               | 10.1038/s41591-021-01292-y    | UK/US/Sweden    | Original research (brief communication/letter) | Cohort study                     | yes; country, sex, age (years), age group, obese (%), BMI, comorbidities, IMD, hospital visits, symptoms |
| Nature Medicine         | 15/04/2021   | Development of a human skin commensal microbe for bacteriotherapy of atopic dermatitis and use in a phase 1 randomized clinical trial | 10.1038/s41591-021-01256-2    | US              | Original research (full paper) | RCT                              | yes; age, sex, ethnicity and race                                      |
| Nature Medicine         | 15/04/2021   | Fetal cranial growth trajectories are associated with growth and neurodevelopment at 2 years of age: INTERBIO-21st Fetal Study | 10.1038/s41591-021-01280-2    | Global          | Original research (brief communication/letter) | Cohort study                     | yes; Sex, SES (university education, married/living as married, work outside of home), health status outcomes |
| Nature Medicine         | 15/04/2021   | altered perivascular fibroblast activity predicts ALS disease onset   | 10.1038/s41591-021-01295-9     | Europe          | Original research (brief communication/letter) | Interventional other (lab)      | yes; age, gender                                                        |
| Nature Medicine         | 15/04/2021   | Homozygous BCMA gene deletion in response to anti-BCMA CAR T cells in a patient with multiple myeloma | 10.1038/s41591-021-01245-5     | Germany         | Original research (brief communication/letter) | Other Observational (lab)       | yes; age, sex                                                          |
| Nature Medicine         | 15/03/2021   | Impaired meningeal lymphatic drainage in patients with idiopathic Parkinson’s disease | 10.1038/s41591-020-01198-1     | US/China/USA    | Original research (brief communication/letter) | Case control study              | yes; n (%) female, age,                                                |
| Nature Medicine         | 15/03/2021   | TCR-engineered T cells targeting E7 for patients with metastatic HPV-associated epithelial cancers | 10.1038/s41591-020-01225-1     | US              | Original research (brief communication/letter) | Interventional trial (not RCT)  | yes; age, sex male/female,                                               |
| PLOS One                | 19/04/2021   | Effect of dietary treatment and fluid intake on the prevention of recurrent calcium stones and changes in urine composition: A meta-analysis and systematic review | 10.1371/journal.pone.0250257 | USA/UK/China    | Original research (full paper) | Systematic review and meta-analysis | Yes: Age, male (n)                                                      |
| PLOS One                | 19/04/2021   | Prognostic value of the postoperative neutrophil- | 10.1371/journal.pone.0250091  | USA/UK/China    | Original research (full paper) | Meta-analysis                     | Yes: age group, male vs female, disease characteristics             |
| Journal | Date          | Title                                                                 | DOI                                      | Country | Region | Study Type      | Design Details                                                                 | Yes: Variables                                                                                           |
|---------|--------------|----------------------------------------------------------------------|-------------------------------------------|---------|--------|----------------|--------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------|
| PLOS One | 19/04/2021   | Predicting poor outcomes in children aged 1–12 with respiratory tract infections: A systematic review | 10.1371/journal.pone.0249533            | USA/UK  | UK     | Original research (full paper) | systematic review (no meta-analysis)                                                                 | Yes: age only                                                                                           |
| PLOS One | 16/04/2021   | Effect of smoking status and programmed death-ligand 1 expression on the microenvironment and malignant transformation of oral leukoplakia: A retrospective cohort study | 10.1371/journal.pone.0250359            | USA/UK  | Japan   | Original research (full paper) | Cohort study                                                                                           | Yes: sex (Male, Female), Age, alcohol drinking, lesion site, disease specific features                   |
| PLOS One | 16/04/2021   | A dose-dependent beneficial effect of methotrexate on the risk of interstitial lung disease in rheumatoid arthritis patients | 10.1371/journal.pone.0250339            | USA/UK  | Poland  | Original research (full paper) | Cohort study                                                                                           | Yes: Age, male sex, disease specific factors                                                                 |
| PLOS One | 16/04/2021   | CT-based determination of excessive visceral adipose tissue is associated with an impaired survival in critically ill patients | 10.1371/journal.pone.0250321            | USA/UK  | Germany | Original research (full paper) | Cohort study                                                                                           | Yes: Gender (male, female), Age, BMI, disease specific features and comorbidities                           |
| PLOS One | 16/04/2021   | Parental educational level and childhood wheezing and asthma: A prospective cohort study from the Japan Environment and Children's Study (plos.org) | 10.1371/journal.pone.0250255            | USA/UK  | Japan   | Original research (full paper) | Cohort study                                                                                           | Yes: Sex (boy/Girl), Child age, mothers educational level, fathers educational level                      |
| PLOS One | 16/04/2021   | The processing of intimately familiar and unfamiliar voices: Specific neural responses of speaker recognition and identification | 10.1371/journal.pone.0250214            | USA/UK  | Canada  | Original research (full paper) | Cohort study                                                                                           | Yes: '8 females', age                                                                                     |
| PLOS One | 16/04/2021   | Pathological complete response of adding targeted therapy to neoadjuvant chemotherapy for inflammatory breast cancer: A systematic review | 10.1371/journal.pone.0250057            | USA/UK  | Global  | Original research (full paper) | systematic review (no meta-analysis)                                                                | Yes: Age, (all females)                                                                                  |
| PLOS One | 16/04/2021   | Dose-response relationships of intestinal organs and excessive mucus discharge after gynaecological radiotherapy | 10.1371/journal.pone.0250004            | USA/UK  | Sweden  | Original research (full paper) | Cohort study                                                                                           | Yes: Age, (all females)                                                                                  |
| BMJ     | 14/04/2021   | Associations of healthy lifestyle and socioeconomic status with mortality and incident cardiovascular disease: two prospective cohort studies | 10.1136/bmj.n604                     | UK      | USA/UK  | Original research (full paper) | Cohort study                                                                                           | Yes: mean age, men, white ethnicity or race, married, household income, occupation, education, health insurance, socio-economic index, |
| BMJ | Date       | Title                                                                 | DOI                      | Country                   | Design                           | Description                                                                 |
|-----|------------|----------------------------------------------------------------------|--------------------------|---------------------------|----------------------------------|----------------------------------------------------------------------------|
| BMJ | 14/04/2021 | Continued versus discontinued oxytocin stimulation in the active phase of labour (CONDISOX): double blind randomised controlled trial | 10.1136/bmj.n716         | UK Denmark/Netherlands    | Original research (full paper)   | smoking, alcohol, diet, BMI, comorbidities                                |
| BMJ | 07/04/2021 | Linked electronic health records for research on a nationwide cohort of more than 54 million people in England: data resource | 10.1136/bmj.n826         | UK UK                     | Original research (full paper)   | yes; age, sex, ethnicity, comorbidities                                   |
| BMJ | 06/04/2021 | E-health StandingTall balance exercise for fall prevention in older people: results of a two year randomised controlled trial | 10.1136/bmj.n740         | UK Australia              | Original research (full paper)   | yes; gender, age, dependant child in household, clinical vulnerability, household member with chronic illness, employment status, socioeconomic grade, index of multiple deprivation, highest educational or professional qualification, ethnicity (white, Asian, mixed, Arab or other (don’t know or prefer not to), living alone, marital status, employment, hardship |
| BMJ | 31/03/2021 | Adherence to the test, trace, and isolate system in the UK: results from 37 nationally representative surveys | 10.1136/bmj.n608         | UK UK                     | Original research (full paper)   | yes; age, (all women)                                                    |
| BMJ | 31/03/2021 | Post-covid syndrome in individuals admitted to hospital with covid-19: retrospective cohort study | 10.1136/bmj.n693         | UK UK                     | Original research (full paper)   | yes; age, sex (men/women), ethnicity (white, Asian, mixed/other, unknown) index of multiple deprivation category |
| BMJ | 24/03/2021 | Comparative efficacy of interventions for reducing symptoms of depression in people with dementia: systematic review and network meta-analysis | 10.1136/bmj.n532         | UK Canada                 | Original research (full paper)   | systematic review and meta-analysis                                     |
| BMJ | 24/03/2022 | Association of spontaneous abortion with all cause and cause specific premature mortality: prospective cohort study | 10.1136/bmj.n530         | US                        | Original research (full paper)   | yes; age, (all women) race/ethnicity (n %) non-Hispanic white, non-Hispanic black, Hispanic and other |
| Journal | Date | Title | DOI | Country | Study Type | Details |
|---------|------|-------|-----|---------|------------|---------|
| BMJ | 23/02/2022 | Age dependent associations of risk factors with heart failure: pooled population based cohort study | 10.1136/bmj.n461 | UK | Original research (full paper) | Yes: age, male sex, white ethnicity, |
| BMJ | 18/03/2021 | Association between living with children and outcomes from covid-19: OpenSAFELY cohort study of 12 million adults in England | 10.1136/bmj.n628 | UK | Original research (full paper) | Yes: age (groups), female sex, ethnicity (white, mixed, south Asian, black, other), Index of multiple deprivation, over 3 adults in a household, |
| Cochrane | 15/04/2021 | Abdominal ultrasound and alpha-foetoprotein for the diagnosis of hepatocellular carcinoma in adults with chronic liver disease | 10.1002/14651858.CD013346.pub2 | UK | Original research (full paper) | Systematic review and meta-analysis | Yes: age, gender individually |
| Cochrane | 15/04/2021 | Thrombolytic therapy for pulmonary embolism | 10.1002/14651858.CD004437.pub6 | UK | Original research (full paper) | Systematic review and meta-analysis | Yes: Age, ‘sex’ as men and women, when reporting characteristics of studies included |
| Cochrane | 14/04/2021 | Dopamine agonists for preventing ovarian hyperstimulation syndrome | 10.1002/14651858.CD008605.pub4 | UK | Original research (full paper) | Systematic review and meta-analysis | Yes: Age if reported in primary study, all women, |
| Cochrane | 14/04/2021 | Regular treatment with formoterol and an inhaled corticosteroid versus regular treatment with salmeterol and an inhaled corticosteroid for chronic asthma: serious adverse events | 10.1002/14651858.CD007694.pub3 | UK | Original research (full paper) | Systematic review and meta-analysis | Yes: Age, (no sex or gender reported) |
| Cochrane | 14/04/2021 | Botulinum toxin type A versus anticholinergics for cervical dystonia | 10.1002/14651858.CD004312.pub3 | Portugal | Original research (full paper) | Systematic review and meta-analysis | Yes: age, % female |
| Cochrane | 13/04/2021 | Non-steroidal anti-inflammatory drugs (NSAIDs) for trigger finger | 10.1002/14651858.CD012789.pub2 | Singapore | Original research (full paper) | Systematic review and meta-analysis | Yes: age, gender (male/female) |
| Cochrane | 12/04/2021 | Monitoring of stimulated cycles in assisted reproduction (IVF and ICSI) | 10.1002/14651858.CD005289.pub4 | UK | Original research (full paper) | Systematic review and meta-analysis | Yes: age, all female, |
| Cochrane | 10/04/2021 | Treatment for bleeding oesophageal varices in people with decompensated liver cirrhosis: a network meta-analysis | 10.1002/14651858.CD013155.pub2 | UK | Original research (full paper) | Systematic review and meta-analysis | Yes: age, ‘females n and %’ |
| Cochrane | 07/04/2021 | Anti-seizure medications for Lennox-Gastaut syndrome | 10.1002/14651858.CD003277.pub4 | Italy | Original research (full paper) | Systematic review and meta-analysis | Yes: Age, sex (as per study), race/ethnicity |
| Cochrane | 06/04/2021 | Primary prevention of variceal bleeding in people with oesophageal varices due to liver cirrhosis: a network meta-analysis | 10.1002/14651858.CD013121.pub2 | UK | Original research (full paper) | Systematic review and meta-analysis | Yes: mean age, females n and % |
| Journal                         | Date          | Title                                                                 | DOI                                      | Country 1 | Country 2 | Article Type                        | Study Type            | Variables               |
|-------------------------------|---------------|----------------------------------------------------------------------|------------------------------------------|-----------|-----------|------------------------------------|-----------------------|-------------------------|
| Cell Metabolism               | 06/04/2021    | Hyocholic acid species improve glucose homeostasis through a distinct TGR5 and FXR signaling mechanism | 10.1016/j.cmet.2020.11.017              | UK        | China     | Original research (full paper)     | Cohort study          | Yes, Age, sex           |
| Cell Metabolism               | 02/03/2021    | The pyruvate-lactate axis modulates cardiac hypertrophy and heart failure | 10.1016/j.cmet.2020.12.003              | UK        | USA       | Original research (full paper)     | Other Observational (lab) | Yes, Age, sex           |
| Cell Metabolism               | 02/02/2021    | Neutrophils Fuel Effective Immune Responses through Gluconeogenesis and Glycogenesis | 10.1016/j.cmet.2020.11.016              | UK        | Scotland  | Original research (full paper)     | Other Observational (lab) | Yes, Age, sex           |
| Cell Metabolism               | 05/01/2021    | Acetyl-CoA Synthetase 2: A Critical Linkage in Obesity-Induced Tumorigenesis in Myeloma | 10.1016/j.cmet.2020.12.011              | UK        | USA       | Original research (full paper)     | Other Observational (lab) | Yes, age, sex           |
| Cell Metabolism               | 01/12/2020    | Succinyl-CoA Ligase Deficiency in Pro-inflammatory and Tissue-Invasive T Cells | 10.1016/j.cmet.2020.10.025              | UK        | USA       | Original research (full paper)     | Other Observational (lab) | Yes, age, sex           |
| Cell Metabolism               | 01/12/2020    | SARS-CoV-2 Cell Entry Factors ACE2 and TMPRSS2 Are Expressed in the Microvasculature and Ducts of Human Pancreas but Are Not Enriched in β Cells | 10.1016/j.cmet.2020.11.006              | UK        | USA       | Original research (full paper)     | Other Observational (lab) | Yes, age, sex, ethnicity, BMI |
| Cell Metabolism               | 01/12/2020    | Expression of SARS-CoV-2 Entry Factors in the Pancreas of Normal Organ Donors and Individuals with COVID-19 | 10.1016/j.cmet.2020.11.005              | UK        | USA       | Original research (full paper)     | Case control          | Yes, age, sex, ethnicity, BMI |
| Cell Metabolism               | 01/12/2020    | Elevation of JAML Promotes Diabetic Kidney Disease by Modulating Podocyte Lipid Metabolism | 10.1016/j.cmet.2020.10.019              | UK        | China     | Original research (full paper)     | Other Observational (lab) | Yes, age, sex           |
| Cell Metabolism               | 03/11/2020    | Pyruvate Kinase Controls Signal Strength in the Insulin Secretory Pathway | 10.1016/j.cmet.2020.10.007              | UK        | USA       | Original research (full paper)     | Other Observational (lab) | Yes, age, sex           |
| Cell Metabolism               | 03/11/2020    | Bone Marrow Mesenchymal Stem Cells Support Acute Myeloid Leukemia Bioenergetics and Enhance Antioxidant Defense and Escape from Chemotherapy | 10.1016/j.cmet.2020.09.001              | UK        | Switzerland | Original research (full paper)     | Other Observational (lab) | Yes, age, sex           |
| Science translational medicine| 14/04/2021    | Imaging Enterobacteriales infections in patients using pathogen-specific positron emission tomography | 10.1126/scitranslmed.abe9805            | USA       | USA       | Original research (full paper)     | Other Observational (lab) | Yes: Age, Sex (M/F), weight, medical conditions |
| Science translational medicine| 14/04/2021    | Rituximab-resistant splenic memory B cells and newly engaged naive B cells fuel relapses in patients with immune thrombocytopenia | 10.1126/scitranslmed.abc3961            | USA       | France     | Original research (full paper)     | Other Observational (lab) | Yes: Age, gender (M/F)  |
| Journal                      | Date       | Title                                                                 | DOI                                      | Country   | Country  | Study Type        | Outcome Measures                                                                 |
|------------------------------|------------|-----------------------------------------------------------------------|------------------------------------------|-----------|-----------|------------------|----------------------------------------------------------------------------------|
| Science translational medicine | 07/04/2021 | SerpinB13 antibodies promote β cell development and resistance to type 1 diabetes | 10.1126/scitranslmed.abf1587            | USA       | USA       | Original research (full paper) | cohort study: Yes male to female ratio, age, diagnosis                           |
| Science translational medicine | 07/04/2021 | A selective HDAC8 inhibitor potentiates antitumor immunity and efficacy of immune checkpoint blockade in hepatocellular carcinoma | 10.1126/scitranslmed.aaz6804            | USA       | Hong Kong | Original research (full paper) | Interventional other (Lab): Yes: sex male/female; Age; disease characteristics (in suppl table s1) |
| Science translational medicine | 07/04/2021 | Urolithin A improves muscle function by inducing mitophagy in muscular dystrophy | 10.1126/scitranslmed.abb0319            | USA       | Switzerland | Original research (full paper) | Interventional other (Lab): Yes (under 'Human Cells' heading) age, sex male (sex linked disorder), |
| Science translational medicine | 31/03/2021 | Soft, skin-interfaced sweat stickers for cystic fibrosis diagnosis and management | 10.1126/scitranslmed.abd8109            | USA       | USA       | Original research (full paper) | Interventional other (Lab): Yes Age, gender Female/male                           |
| Science translational medicine | 31/03/2021 | Clearance of pegylated interferon by Kupffer cells limits NK cell activation and therapy response of patients with HBV infection | 10.1126/scitranslmed.aba6322            | USA       | USA       | Original research (full paper) | cohort study: Yes (supp tab s1): sex %Male, % female, race 'Asian, Black, Caucasian', BMI, disease characteristics, |
| Science translational medicine | 31/03/2021 | Increasing breast milk betaine modulates Akkermansia abundance in mammalian neonates and improves long-term metabolic health | 10.1126/scitranslmed.abb0322            | USA       | Spain     | Original research (full paper) | cohort study: Yes: Age, gender (M/F)                                             |
| Science translational medicine | 31/03/2021 | Transcriptional networks in at-risk individuals identify signatures of type 1 diabetes progression | 10.1126/scitranslmed.abd5666            | USA       | UK        | Original research (full paper) | cohort study: Yes: Age, race, race-ethnicity                                       |
| Science translational medicine | 17/03/2021 | GDE2-RECK controls ADAM10 α-secretase–mediated cleavage of amyloid precursor protein | 10.1126/scitranslmed.abe6178            | USA       | USA       | Original research (full paper) | Interventional other (Lab): Yes: age, gender male/female, Race (White, Black)     |
### Table 2: Excluded articles from each journal

| Journal            | Excluded Articles                                                                 |
|--------------------|-----------------------------------------------------------------------------------|
| NEJM               | Non-research articles (n= 24); Animal studies/Other non-human (n=0); human research with no participant level data reported (n=2) |
| Lancet            | Non-research articles (n= 82); Lancet Animal studies/Other non-human (n=0); human research with no participant level data reported (n=2) |
| JAMA              | Non-research articles (n=51); Animal studies/Other non-human (n=0); No participant level data reported (n=2) |
| PNAS              | Non-research articles (n=69); Animal studies/Other non-human (n=30); No participant level data reported (n=4) |
| Nature Medicine   | Non-research articles (total n=18); Animal studies/Other non-human (n=1); No participant level data reported (n=7) |
| PLOSOne           | Non-research articles (n=0); Animal studies/Other non-human (n=4); No participant level data reported (n=0) |
| BMJ               | Non-research articles (n=141); Animal studies/Other non-human (n=1); No participant level data reported (n=2) |
| Cochrane          | Non-research papers (n=0); Animal studies/Other non-human (n=0); No participant level data reported (n=1, but only because no studies included) |
| Cell metabolism   | Non-research articles (n=54); Animal studies/Other non-human (n=33); No participant level data reported (n=0) |
| Science Translational Medicine | Non-research articles (n=3); Animal studies/Other non-human (n=14); No participant level data reported (n=5) |
Table 3: Terms accepted within papers for reporting gender, ethnicity or SES

| Accepted gender reporting terms | Accepted ethnicity reporting terms | Accepted Socio-economic status reporting terms |
|---------------------------------|------------------------------------|-----------------------------------------------|
| Male (number and/or %)          | Race                               | Employment status                              |
| Female (number and/or %)        | Ethnicity                          | University level education                     |
| Gender: Male/Female (number and/or %) | Race or ethnic group               | Urban residence                                |
| Sex: Male/Female (number and/or %) | Race/ethnicity                  | Index of multiple deprivation                 |
| Male/ female /prefer not to say | Population sector                  | Region of England                              |
| Male/ Female/ other             | Geographic region                  | Education                                      |
| Male: Female ratio              | Results by continent               | Health Insurance (private/semiprivate,         |
| All female sex                  | Geographic location                | government subsidised, uninsured               |
| All Male sex                    | Geographic region                  | Number of people in household                 |
| Boy/Girl                        | Race or ethnic group               | Educational attainment                         |
| Gender                          | Ethnicity                          | Poverty Index ratio                            |
| Sex                             | Non-US born                        | Mothers educational level/ Fathers educational level |
| M/F                             | Native to America                  | Over 3 adults in household                     |
|                                 | Native American                    | Employment status                              |
|                                 | White race                         | Hardship                                       |
|                                 | Geographic origin                  | SES (university education, married/living as married, work outside of home), |
|                                 |                                   | Household income                               |
|                                 |                                   | Socioeconomic grade                            |
|                                 |                                   | Highest educational or professional qualification |
|                                 |                                   | Socioeconomic Index                            |
# PRISMA 2020 Checklist

| Section and Topic | Item # | Checklist item | Location where item is reported |
|-------------------|--------|----------------|---------------------------------|
| **TITLE**         |        |                |                                 |
| Title             | 1      | Identify the report as a systematic review. | 1 ('targeted' as we have adapted our approach) |
| **ABSTRACT**      |        |                |                                 |
| Abstract          | 2      | See the PRISMA 2020 for Abstracts checklist. | As per BMJ Open |
| **INTRODUCTION**  |        |                |                                 |
| Rationale         | 3      | Describe the rationale for the review in the context of existing knowledge. | 4 |
| Objectives        | 4      | Provide an explicit statement of the objective(s) or question(s) the review addresses. | 4 |
| **METHODS**       |        |                |                                 |
| Eligibility criteria | 5 | Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses. | 5 |
| Information sources | 6 | Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted. | 4-5 |
| Search strategy   | 7      | Present the full search strategies for all databases, registers and websites, including any filters and limits used. | 4-5 |
| Selection strategy | 8 | Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process. | 5 |
| Data collection process | 9 | Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process. | 4-5 |
| Data items        | 10a    | List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect. | 5 |
|                   | 10b    | List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information. | 5 |
| Study risk of bias assessment | 11 | Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process. | n/a |
| Effect measures   | 12     | Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results. | 5-6 |
| Synthesis methods | 13a    | Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)). | 5-6 |
|                   | 13b    | Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions. | 5-6 |
|                   | 13c    | Describe any methods used to tabulate or visually display results of individual studies and syntheses. | 5-6 |
|                   | 13d    | Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used. | 5-6 |
|                   | 13e    | Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression). | 5-6 |
|                   | 13f    | Describe any sensitivity analyses conducted to assess robustness of the synthesized results. | n/a |
# PRISMA 2020 Checklist

| Section and Topic            | Item # | Checklist item                                                                 | Location where item is reported |
|------------------------------|--------|-------------------------------------------------------------------------------|---------------------------------|
| Reporting bias assessment    | 14     | Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases). | n/a                             |
| Certainty assessment         | 15     | Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome. | n/a                             |
| RESULTS                      |        |                                                                               |                                  |
| Study selection              | 16a    | Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram. | 6                               |
|                              | 16b    | Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded. | 5-6                             |
| Study characteristics        | 17     | Cite each included study and present its characteristics.                     | Online suppl.                   |
| Risk of bias in studies      | 18     | Present assessments of risk of bias for each included study.                  | n/a                             |
| Results of individual studies| 19     | For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots. | n/a                             |
| Results of syntheses         | 20a    | For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies. | Online suppl.                   |
|                              | 20b    | Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect. | 5-6                             |
|                              | 20c    | Present results of all investigations of possible causes of heterogeneity among study results. | 5-6                             |
|                              | 20d    | Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results. | 5-6                             |
| Reporting biases             | 21     | Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed. | n/a                             |
| Certainty of evidence        | 22     | Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed. | n/a                             |
| DISCUSSION                   |        |                                                                               |                                  |
| Discussion                   | 23a    | Provide a general interpretation of the results in the context of other evidence. | 7-8                             |
|                              | 23b    | Discuss any limitations of the evidence included in the review.               | 7-8                             |
|                              | 23c    | Discuss any limitations of the review processes used.                        | 7-8                             |
|                              | 23d    | Discuss implications of the results for practice, policy, and future research. | 7-9                             |
| OTHER INFORMATION            |        |                                                                               |                                  |
| Registration and protocol    | 24a    | Provide registration information for the review, including register name and registration number, or state that the review was not registered. | Not registered                  |
|                              | 24b    | Indicate where the review protocol can be accessed, or state that a protocol was not prepared. | Agreed between authors but not made available publicly. |
## PRISMA 2020 Checklist

| Section and Topic | Item # | Checklist item                                                                                                                                                                                                 | Location where item is reported |
|-------------------|--------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------|
| 24c               | Describe and explain any amendments to information provided at registration or in the protocol.                                                                                                               | n/a                              |
| Support           | 25     | Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.                                                                               | 9                                |
| Competing interests | 26     | Declare any competing interests of review authors.                                                                                                                                                           | 9                                |
| Availability of data, code and other materials | 27     | Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review. | Online Suppl.                    |

*From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71

For more information, visit: [http://www.prisma-statement.org/](http://www.prisma-statement.org/)*