**Article:**
Uttley, L., Weng, Y. and Falzon, L. (2026) What's the meta now? More updates on the problems with systematic reviews. Journal of Clinical Epidemiology. 112393. ISSN: 0895-4356

https://doi.org/10.1016/j.jclinepi.2026.112393

# Title: What's the Meta Now? More Updates on the Problems with Systematic Reviews

**Authors: Lesley Uttley[1*], Yuliang Weng[2], Louise Falzon[1].**

[1] School of Medicine and Population Health, University of Sheffield, United Kingdom.

[2] IT Services, University of Sheffield, United Kingdom.

## Abstract

**Background:** Systematic reviews are intended to provide trustworthy evidence synthesis, yet previous iterations of this living review have identified numerous recurring problems in their conduct and reporting. This article presents the third version and second update of the living systematic review examining issues raised across the academic literature.

**Methods:** Using consistent eligibility criteria and methods from earlier versions, literature searches were updated to May 2025. Eligible meta-research and editorial articles describing problems with systematic reviews were analysed to identify emerging themes. Additionally, four basic indicators of methodological quality of the included meta-research were presented across review versions.

**Results:** The update included 209 additional articles. Critically low methodological quality and absence of protocols remained among the most frequently reported issues in systematic reviews across disciplines and journals but notably in evidence underpinning clinical practice guidelines. Spin in abstracts and conflicts of interest continued to be common. Apparent improvements in reporting quality were inconsistent, with modest gains in some full-text reporting but persistent deficiencies in abstracts. Authorship diversity of systematic reviews improved in gender representation but remained geographically concentrated in high-income countries, and primary research included in reviews similarly lacked global representativeness. The issue of misalignment between systematic review evidence bases and global burden of disease bring the total number of problems with systematic reviews to 69. Emerging use of automation and artificial intelligence was variably reported. Descriptive comparison of meta-research articles over the three versions of this living review suggests a greater proportion meeting basic quality indicators in more recent updates.

**Conclusions:** Across successive updates, problems with systematic reviews remain widespread and consistent rather than isolated. Incremental reporting improvements coexist

with persistent concerns about transparency, bias, and representativeness. Future efforts should prioritise evaluating interventions and aligning research incentives to support genuinely trustworthy evidence synthesis.

*Keywords*: Systematic review; research integrity; trustworthy evidence; global research representation; authorship; conflicts of interest; meta-research.

## What is new?

Key findings

- Analysis of 209 newly included articles indicates critically low methodological quality and absent protocols in published systematic reviews, including those underpinning major clinical practice guidelines.
- Geographic representation of systematic review evidence is concentrated in high-income countries, limiting the global applicability of evidence syntheses.
- Emerging automation and artificial intelligence tools offer potential solutions to scale and efficiency challenges but are themselves inconsistently reported and evaluated

What this adds to what was known?

- Although some modest gains in reporting standards for systematic reviews are observed, improvements are uneven and often confined to full-text reporting, while abstract reporting remains limited.
- Spin in abstracts and underreported financial conflicts of interest continue to distort interpretation of systematic reviews, with evidence that commercial ties can influence design choices and outcome emphasis.
- Misalignment between systematic review evidence bases and global burden of disease brings the total number of problems with systematic reviews to 69.

What is the implication and what should change now?

- Meta-scientific efforts should redirect attention towards rigorous assessment of how methodological weaknesses affect validity of conclusions and clinical decision-making beyond repeatedly cataloguing known deficiencies.
- Journals, peer reviewers and guideline developers should enforce mandatory protocol registration, detailed methods reporting, and full conflict-of-interest disclosure for systematic review authors.

## Introduction

Well conducted systematic reviews are the benchmark for trustworthy evidence synthesis. Previously, sixty-eight discrete problems with published systematic reviews were identified through 637 meta-research and editorial academic articles from a living systematic review ("Systematic Reviewlution" www.systematicreviewlution.com). (1,2) This article presents the second update, and third version, of a living systematic review. The aim is to describe emerging themes from systematic review problems identified from newly published literature in the annual living review update.

## Methods

This living review follows a prospectively registered protocol (PROSPERO ID: CRD42020181371; OSF: https://osf.io/2hmv9/) and the full methods are described in the first published version.(1) Briefly, the review includes empirical meta-research studies (e.g. cross-sectional, meta-epidemiological, surveys) and non-empirical articles (e.g. letters and discussion pieces) that explicitly report problems in the conduct or reporting of published systematic reviews. Comprehensive searches of bibliographic databases and supplementary citation searching were amended including cited reference searching using the currently included studies. Update searches were run on 7th and 13th May 2025, covering the period from May 2023 to May 2025. Records were screened independently in duplicate, with disagreements resolved by consensus. Data were extracted using a structured, purpose-built database and verified by a second reviewer. Extracted data included study characteristics, types of problems identified, and whether findings related to methodological quality, reporting, or interpretation. Study quality indicators were pragmatically assessed using key methodological features (e.g. protocol availability, sampling strategy, duplicate processes, and reproducibility of methods). Findings were synthesised using inductive thematic analysis, with problems categorised into overarching domains of systematic review conduct.

The outcomes of interest in this living review update were:

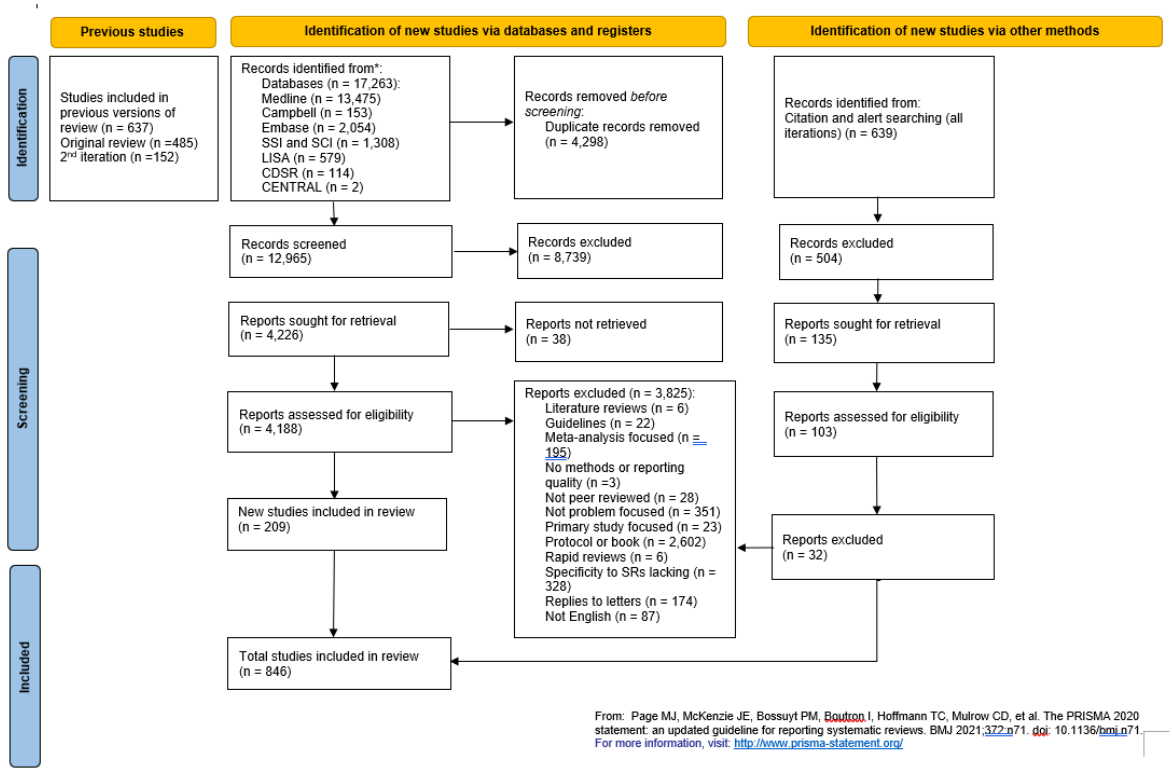
1) Themes from recent prevalent problems in systematic review conduct or reporting identified from eligible articles in the update searches.

2) To assess the quality of included meta-research articles across the three versions of the review.

## Results

Literature searches yielded 5,515 citations after duplicates were removed. Following screening of titles and abstracts 2,677 full texts were assessed for eligibility, and 209 of these were deemed as eligible for inclusion into the living systematic review (see Figure 1).

*Figure 1: PRISMA Flow diagram of identification and screening of articles included in the second living systematic review update*




Recurring Quality Issues & Biases in Evidence Underpinning Clinical Practice Guidelines

The majority of included meta-research papers highlight critically low or low methodological quality (mostly assessed according to AMSTAR-2 guidelines), among the majority of systematic reviews analysed (115/209). Lack of registered or publicly available protocol is the second most commonly occurring problem with systematic reviews (81/209). As in previous updates, these recurring findings have been noted extensively across clinical specialities, journals and fields. (2)

Five included meta-research articles highlight poor methodological and/or reporting conduct or systematic reviews underpinning clinical practice guidelines in the US for cancer network

guidelines, (3) US cutaneous melanoma, (4) US tobacco use disorder, (5) and US alcohol use disorder, (6) and American dietary guidelines, (7). One article highlights the same low methodological conduct supporting UK vascular surgery guidelines. (8)

Spin, also termed ‘narrative bias’, continues to be a prevalent issue where results are distorted in titles and abstracts to suggest more favourable conclusions than the data support. In research on cannabinoids for pain, approximately 17% of systematic review abstracts exhibited moderate to severe bias, consistently favouring the experimental treatment's efficacy and safety even when supportive data were inadequate.(9) Similarly, in orthodontic literature, exaggerated claims appear in 40.8% to 64.4% of systematic review abstracts, primarily manifesting as misleading reporting that omits adverse effects, selective reporting of beneficial outcomes, or misleading interpretations that ignore statistical non-significance.(10,11) Because clinicians often rely solely on abstracts due to time constraints, paywalls, or language barriers, this trend is particularly concerning, especially as this study shows that the presence of spin was often not significantly associated with funding sources, publication year, or the number of authors.(11) Consequently, readers can be misled by over-statements in abstracts that may not faithfully reflect the full body of evidence.

Evidence on conflicts of interest (COI) from long-acting injectable antipsychotics demonstrates that researchers with commercial ties often make ‘industry-friendly’ choices, such as selecting weak pre-post study designs to inflate perceived benefits, misrepresenting harm data in the main text of papers, and prioritising surrogate markers over patient-centred outcomes like quality of life. (12) Furthermore, a cross-sectional study on ADHD treatments found that COIs were underreported in 18% of reviews compared to publicly available data, with approximately half of all included reviews containing at least one author with a financial tie. (13) While over 80% of these reviews were rated as having a high RoB, the sources highlight a complex trend where authors of lower-bias reviews are paradoxically more likely to disclose their financial interests than those producing higher-bias work, underscoring the urgent need for stricter independent standards to prevent commercial interests from "co-opting" medical knowledge.

Are systematic reviews improving over time?

A number of observational meta-research studies examine characteristics of review conduct over time, with mixed conclusions. Some exploratory studies highlight modest improvements in conduct assessed through adherence to methodological and reporting guidelines for 215 US systematic reviews (8) for 1176 systematic reviews published in Chinese journals between 2009 and 2022 (14) and 153 systematic reviews about clinical pharmacy services. (15) However, one meta-epidemiological study found critically low methodological quality in 90% of 100 systematic reviews of physiotherapy for musculoskeletal conditions did not improve

over a ten-year period (2012–2022). (16) Moreover, reporting quality of full-text reports was not found to correspond with reporting quality of abstracts. A five-year follow-up study (comparing 2019 to 2024) found modest improvements in full-text reporting (adherence to PRISMA-DTA rose from 71% to 78%). (17) However, abstract reporting remained stagnant at approximately 52% adherence. Adherence to Cochrane guidance for incorporating non-randomised studies of interventions (NRSI) was also reported to remain generally low. Although slight improvement was seen between 2019 and 2023 (such as increased use of the recommended ROBINS-I risk of bias tool) fewer than half of the reviews justified why they included NRSI. (18) For Cochrane reviews, reporting quality in abstracts (measured by PRISMA-A) initially improved between 2000 and 2015 but remained unchanged between 2016 and 2022, despite abstracts becoming significantly longer. (19) For Cochrane abstracts, reporting of the phrase 'standard Cochrane methods' has doubled (from 27.4% in 2016 to 53.2% in 2023), and the reporting of GRADE assessments has more than doubled (from 30.7% to 74.2%). (20) However, this article finds that relying on the 'standard methods' phrase often results in authors reporting fewer specific methodological details and reporting more information on results.

Reporting COIs increased from 79% to 94% from 746 systematic reviews indexed on PubMed between 2010 and 2019, however the actual prevalence of individual financial COIs remained steady at approximately 22–23% (when excluding reviews from China).(21) This study highlights a surge of research from China, which grew from 4% to 21% of the total sample and notes that China reported zero individual financial COIs, contrasting sharply with the 50% COI rate found in Cochrane reviews for the same period.

Does Methodological Conduct Affect the Validity of Conclusions?

While we can claim that poor methodological conduct could compromise the reliability of conclusions, it is not certain whether it materially alters the validity of the conclusions themselves. One study of evidence syntheses on rapid antigen tests during the COVID-19 pandemic, found that despite the poor methodological rigour and high risk of bias observed in most of the included reviews, the diagnostic accuracy estimates were consistent among each other. (22)

In addition to the high frequency of critically low methodological quality observed systematic reviews of oncology, reviews also frequently fail to provide "stop statements" or conclusiveness criteria, which are vital for notifying the scientific community when further research on a topic is unnecessary and would result in a waste of resources. Authors of two articles criticise the included reviews for ignoring opportunity costs, such as whether the high

expense of a routinely recommended scan could be better spent on direct patient care like housing or medication. (23,24)

Trends in Review Authors and Team Composition

Continuing scrutiny in the most recently added problem “lack of diversity in review authorship teams”, reveal some progress in gender diversity alongside a concerning stagnation or decline in geographical and economic representation. Female representation has risen substantially over the last two decades in Cochrane reviews (CDSR), from 15.0% in 1996 to 55.6% by 2023 for first authorship. (25) A separate analysis estimated a similar rate of 56.2% for female first authors in the Cochrane Database for Systematic Reviews (CSDR).(26) Across general academic medicine, female first authorship was estimated to have increased by approximately 6.4% every decade.

Despite gains in lead roles, a "leaky pipeline" persists. Women remain significantly underrepresented in senior (last) author positions, which are often perceived as leadership roles. (27,28) For instance, in oncology CDSR publications, women represented 52.3% of first authors but only 39.4% of last authors. (28) In the fields of psychology and psychiatry there was a higher female representation, with 40% first author and 36.7% last author position, whereas medicine was 25.9% and 19.5%, respectively. The rate of change for psychology and psychiatry female authors was also higher every 10 years: 8.56% (95% CI 6.44-10.6). Absolute increase in women authors was found to be largest for first authors (17%) and smallest for senior authors (9.1%). (26)

High income country (HIC) dominance over low- and middle-income countries (LMICs) authorship

Gains in gender balance are tempered by limited geographical representation. Authorship remains heavily concentrated in high-income, English-speaking countries, particularly the United Kingdom, Australia, and the United States**.** (25) LMIC authors constitute less than 2% of the total, with the UK (37.8%) and Canada (10.9%) leading representation. (29) Similarly, 90.5% of authors for glaucoma reviews are from HICs, primarily based in Europe (53.8%) and the Americas (34.8%), with no authors affiliated with institutions in Africa or low-income countries. (30) Representation from authors in LMIC authors within Cochrane reviews has declined recently, dropping from a peak of 23.2% in 2012 to 18.4% in 2023. Non-Cochrane systematic reviews demonstrate much higher diversity, with 50.8% of first authors from LMICs in 2023. This trend is largely driven by the surge in evidence synthesis production from emerging economies, particularly China.(21,25) While authorship from non-English speaking

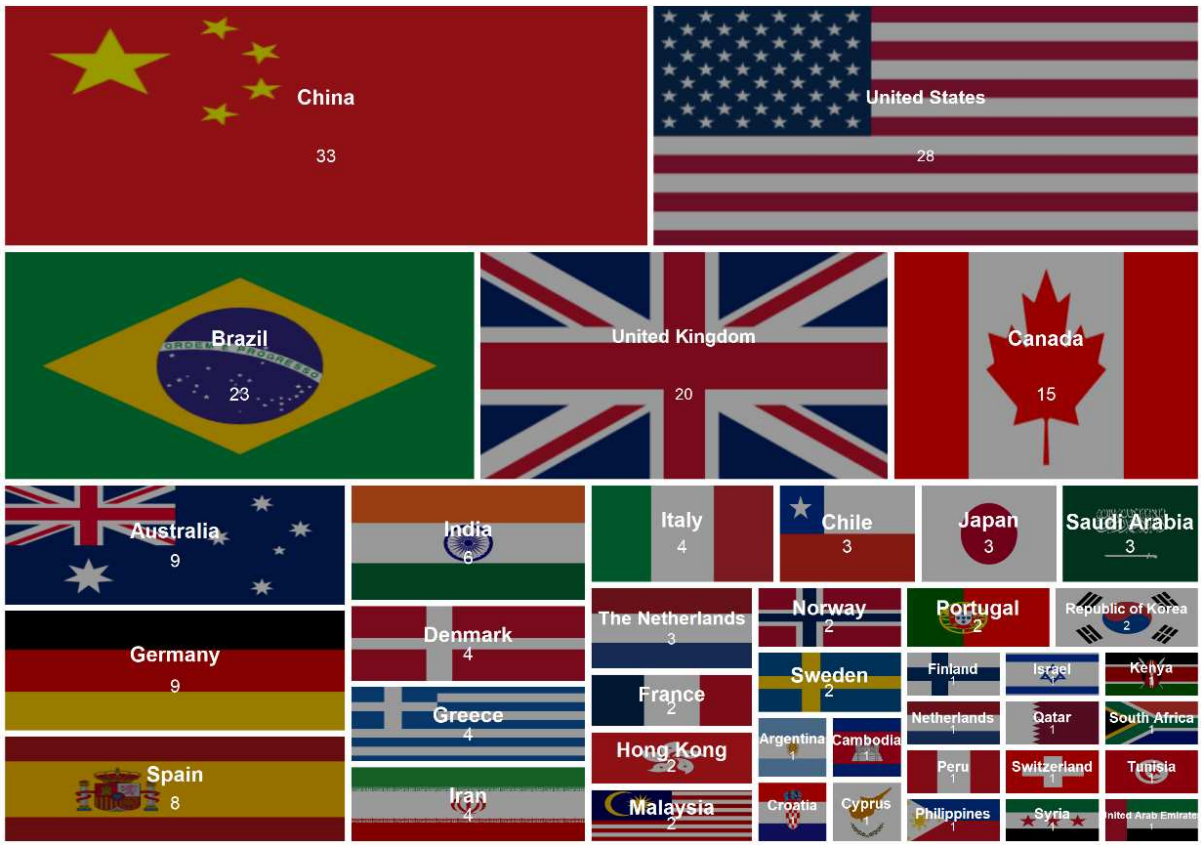
countries has grown, representation from Spanish-speaking authors remains low (approximately 3–4%), despite the presence of dedicated regional networks.

HIC Dominance over LMIC Research

Systematic reviews feature a heavy concentration of primary research in HICs at the expense LMICs. Systematic reviews predominantly synthesise data from high-income regions. In glaucoma research, 72.1% of primary studies were conducted in HICs, mostly in the Americas and Europe. (30) For interpersonal violence interventions, 71% of systematic reviews focused on HIC settings, while only 12% addressed LMICs. (31) Specific types of research, such as school violence interventions, were found to be set exclusively in HICs.

These geographic trends indicate a disconnect where regions with the highest health burdens have the least representation. While the greatest burden of cardiovascular disease and interpersonal violence (such as homicide) is borne by LMICs, these areas are rarely the focus of or contributors to the highest levels of evidence. (29,31) In addition to authorship parasitism where HIC authors monopolise authorship of LMIC research, this lack of parity can result in a "foreign gaze" or highly academic narratives that lack practical relevance for the global majority populations most affected and limits generalisability of systematic reviews. (29,30) This brings the total number of problems with systematic reviews identified by this living review to 69. The problem is with the coverage and representativeness of evidence in evidence syntheses and has therefore been categorised under the overarching domain of “Inclusive/Comprehensive”. Meta-research scrutinising published systematic reviews predominantly originate from HIC as highlighted in the previous update(2) and illustrated in Figure 2.

*Figure 2: Treemap of Countries Affiliated with 209 Included Meta-Research Articles*




Artificial intelligence (AI) as an Emerging Topic

Artificial intelligence is increasingly documented in meta-research of systematic reviews particularly for high-priority health questions in fast moving research areas.(32) Given that this living review is meta upon meta upon meta, there is a natural delay for capturing real-time developments in research practice. As such, an analysis of Campbell reviews found that very few reviews reported the actual use of automation, AI/ML, or text-mining tools in their searches (33) and AI focus is on its diagnostic potential in primary studies rather than its utility in the research workflow itself.(34,35) However, while the AI-assisted diagnostic technology in areas is promisingly efficient, it is noted to require human verification for validation as it has not yet reached the level of reliability seen in LLM-driven text screening guidelines for systematic reviews.(35) This article noted that 7 out of 11 systematic reviews on AI cephalometry presented a high risk of bias, indicating deficiencies in the reporting or conduct of reviews of AI-assisted research. Reporting of artificial intelligence in primary studies benefits from adherence to guidelines to improve such as CONSORT-AI for randomised controlled trials, SPIRIT-AI for protocols and TRIPOD-AI for prediction models. Large Language Model (LLM) abstract screening is emerging as a core component of automated "Evidence Pipelines" designed to manage the exponential growth of primary literature in Cochrane review production systems.(36,37) In terms of reviewing the use of AI in clinical medicine, one study

examining AI in systematic reviews proposes the CLASMOD-AI checklist, highlighting the need for improved reporting on AI metrics in evidence syntheses, particularly regarding training of AI models and dataset quality. (34)

Do Meta-Research Studies Adhere to Basic Methodological Standards?

The high volume of included articles that continue to document flawed systematic reviews over time begs the question: what is the methodological quality of the 846 articles scrutinising systematic reviews? To assess this, four characteristics relevant to methodological and reporting quality extracted in this review are compared across the three living review versions. Details on the full criteria used for data extraction and quality assessment in this review are provided in the supplementary appendix.

Four quality assessment criteria selected to indicate basic methodological standards in empirical meta-research relate to:

1. availability of a study protocol;

2. Whether the sampling strategy for systematic review selection was fair (comprehensive, random, consecutive, purposeful);

3. Whether study selection or data extraction decisions were double checked by a second reviewer, and;

4. Whether the methods of the study are described in enough detail to allow replication (based on provision of searches strategies and references lists of included studies).

All articles including non-empirical articles were eligible as some quality criteria are applicable as “research letters”.

As shown in Table 1, the number of included articles varies between versions, and updates contain some publications from previous years due to indexing delays in bibliographic databases and supplementary searches to identify missed papers (indicated by the mean). The median reflects the central tendency in the temporal distribution of the dataset and the mode highlights clustering in the most recent publication years.

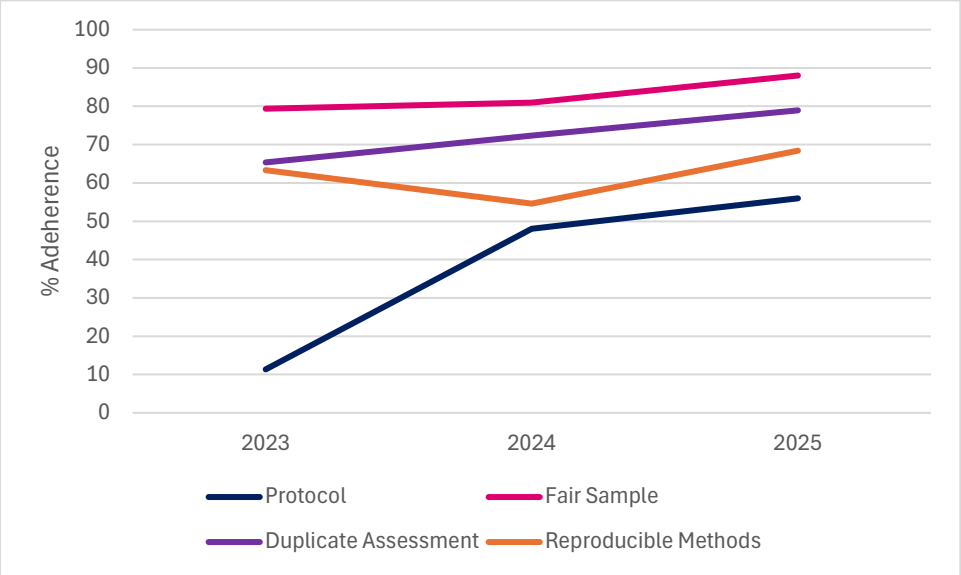
*Table 1: Publication Year of Included Articles Across Living Review Versions*

| Review version | No. Articles | Publication Year | | | |
|---|---|---|---|---|---|
| | | Range | Mean | Median | Mode |
| Original (2023) | 485 | 2000-2021 | 2014 | 2016 | 2019 |
| 1st Update (2024) | 152 | 2020-2024 | 2021 | 2022 | 2022 |

| | | | | | |
|---|---|---|---|---|---|
| 2nd Update: (2025) | 209 | 2019-2025 | 2023 | 2024 | 2024 |

This analysis was not prespecified or powered to detect significant changes. However, despite the differences in article counts between versions, the proportions of meta-research articles shown in Figure 3, meeting basic criteria of quality indicate a positive direction for meta-scientific analysis of systematic reviews. Whilst a promising trend in adherence to basic methodological quality is observed, it is not possible to definitively conclude the quality of included meta-research has increased over time without evidence of statistical significance.

*Figure 3: Adherence to Basic Quality Standards of Articles Scrutinising Systematic Reviews*




## Discussion

The third version of this living review demonstrates an ongoing gap between the intended standards of systematic reviews as the most reliable form of evidence synthesis and the way they are commonly conducted and reported. Across diverse clinical fields, recently included meta-research repeatedly identifies a range of problems but frequently, critically low methodological quality, frequent absence of protocols, and ongoing biases affecting the trustworthiness of systematic reviews. These problems were not confined to niche areas but extended to reviews underpinning major clinical practice guidelines, raising concerns about the evidentiary foundations of recommendations used in routine patient care. Many of the most frequently reported problems (e.g. lack of protocols, low methodological quality) are consistently identified across multiple independent meta-research studies, whereas evidence

for other issues is more limited or emerging. Potential underpinning reasons for the persistent publication of subpar systematic reviews discussed previously, such as the over-reliance on academic papers as a marker of researcher assessment,(38,39) suggest that structural incentives shaping evidence production and reporting need to be addressed to achieve genuinely trustworthy systematic reviews.

Improvements in systematic reviews over time appear modest and inconsistent. While some studies reported apparent incremental gains in reporting standards or guideline adherence, others found no meaningful change across a decade. Improvements in full-text reporting were not mirrored in abstracts, which remain the primary access point for many readers. Conflicts of interest and spin further compromise reliability. Financial ties were common and sometimes underreported, and narrative distortion frequently favoured beneficial interpretations even when statistical support was weak. These findings reinforce concerns that evidence synthesis can amplify, rather than correct, bias present in primary research.

Authorship and geographic representation patterns reveal additional structural limitations in meta-research production. Despite gains in gender representation for authorship, leadership positions and global participation remain uneven. Concentration of authorship is in high income over low- and middle-income countries. Moreover, evidence from primary studies is heavily derived from high-income countries, which limits representation of populations and contexts bearing the greatest global disease burden. Whilst the issue relates to the inclusivity of evidence contained in evidence syntheses there is a wider structural issue relating to research prioritisation and funding which cannot be solved by systematic reviews alone.

Across successive updates, a greater share of meta-research articles fulfilled each of four indicators of methodological quality, suggesting a possible shift toward improved reporting practices, although the analysis was exploratory and should not be interpreted as evidence of a statistically demonstrable trend. We encourage future meta-research to pursue this using a prespecified and adequately powered study design.

Implications

The implications of these findings extend beyond methodological critique. Clinicians and guideline developers should therefore place greater emphasis on critical appraisal of systematic reviews, including protocol availability, risk of bias assessment, and transparency of methods, rather than relying on systematic review status alone as a marker of quality.

As increasing AI tools in evidence synthesis emerge, there are recommendations and guidance in development to improve the conduct and reporting of AI-assistance in systematic reviews such as Responsible use of AI in evidence SynthEsis (RAISE).(40) Analysis of AI in research, from a snapshot of provision in 2025, describes substantial limitations to AI-assisted research according to the accountability, objectivity, rigour, transparency and inclusivity that exemplify the research integrity expected from systematic reviews.(41) Current meta-research provides limited empirical evidence on whether AI improves or compromises systematic review quality, highlighting an important gap for future evaluation.

Journals, peer reviewers, and funders have a role in enforcing minimum standards for systematic reviews and in prioritising interventions that improve review quality, rather than continuing to document known deficiencies. Without such changes, systematic reviews risk further perpetuating biases in the evidence base. We encourage readers to use this living resource to support best practice and informed interpretation. Included meta-research is heavily weighted toward assessing adherence to established methodological and reporting standards, while comparatively less attention has been given to evaluating how these shortcomings affect the validity of conclusions, their influence on clinical decision-making, and the structural factors shaping the production of systematic reviews. For meta-researchers, the accumulating evidence suggests that rather than repeatedly cataloguing known deficiencies, future work should prioritise testing interventions, evaluating impacts, and understanding how emerging developments, including newer trends in automation and artificial intelligence, alter both risks and opportunities within evidence syntheses.

Limitations

This review synthesises findings reported in meta-research studies and does not attempt to reproduce or independently validate included analyses. Classifications of “critically low” or “low” methodological quality reflect those described within included studies, most commonly using AMSTAR-2. Where adherence to reporting or methodological standards (e.g. PRISMA, AMSTAR) was reported as percentage compliance, we applied a pragmatic threshold of <85% to denote less than ideal conduct; this threshold is arbitrary and used for harmonisation across heterogeneous reporting.

Included articles comprise both empirical meta-research and non-empirical publications (e.g. editorials) to enable comprehensive capture of concerns raised in the literature, however, included articles may differ in evidentiary strength. Furthermore, the synthesis is primarily descriptive and does not formally assess or validate the certainty, consistency, or precision of evidence underpinning each identified problem.

The identification of “problems” is dependent on how issues are defined and reported within the included literature and may therefore be influenced by subjective interpretation. In addition, the meta-research literature itself is likely subject to publication and ascertainment bias, whereby known methodological shortcomings are more likely to be investigated and published in areas where concerns are already documented. This living review was designed to identify and synthesise problems reported in the systematic review literature but does not rank their relative severity or impact on the validity and reliability of review conclusions to provide a formal assessment of which problems are most consequential. To support interpretation of the evidence base, detailed study-level extraction and quality assessment data are available through the accompanying living review website (https://systematicreviewlution.com/), including information on whether included studies reported impacts on review conclusions or interpretation. Accordingly, the frequency with which problems are reported should not be interpreted as a measure of their severity or impact on the validity of systematic review conclusions.

## Conclusions

The goal of this living review is to document issues raised with systematic reviews in order to drive meaningful improvements. Recurring deficiencies in methodological conduct, planning, reporting, and interpretation continue to be identified across clinical specialties and time periods. Although incremental improvements are occasionally observed in systematic review quality, these tend to be uneven and limited rather than demonstrable gains in trustworthiness. The persistence of spin, conflicts of interest, limited global representation, and uncertain relationships between methodological rigour and conclusions suggests that improvements in guidance alone are unlikely to resolve the problem. Progress depends on aligning incentives and practices across the research ecosystem so that future systematic reviews represent genuinely trustworthy evidence syntheses.

Funding

This research is funded by a Transition Support Award to Lesley Uttley from the UK Medical Research Council to investigate human influences in systematic reviews (MR/Z504063/1). This research project is registered with the Open Science Framework (https://osf.io/2hmv9/) and PROSPERO (CRD42020181371).

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors used generative AI tools (OpenAI ChatGPT and Google Gemini NotebookLM) used to support language editing and to assist in summarising information from included studies. Tools were used to assist with drafting and organisation of the manuscript and to facilitate review of extracted information; they were not used to generate primary data, perform thematic analysis or replace critical appraisal. All outputs were reviewed, verified, and edited by the authors, who take full responsibility for the accuracy, interpretation, and integrity of the content.

Acknowledgements

The authors wish to thank previous co-authors, collaborators and steering group members that guided the development and conduct of this living systematic review including Paul Montgomery, Dan Quintana, David Moher, Matthew Page, Chris Carroll, Anthea Sutton, Sarah Dawson, Susan Baxter, Penny Whiting, Julia Little, Andrew Booth, Julian Higgins, Hannah Rothstein, Joanne Yaffe, John Stevens and John Ioannidis.

Data availability

The protocol for this study was registered with PROSPERO (CRD42020181371) and posted on Open Science Framework (OSF) (https://osf.io/2hmv9/). Supporting information for the searches and selection of articles in this review are also available at https://osf.io/2hmv9/ and https://www.systematicreviewlution.com.

Declaration of competing interest

This work is funded by the UKRI Medical Research Council with the intention of upholding best practice for systematic reviews but is conducted independently to any methodological or systematic review organisation. The authors posit that there are problems with many published systematic reviews which aligns with the funded research topic but have no other conflicts of interest to declare.

Supplementary data

Supplementary data to this article is provided as an appendix the manuscript.

CRediT author statement

**Lesley Uttley**: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Writing- Original draft preparation, Writing- Reviewing and Editing, Visualization, Supervision, Project administration, Funding acquisition. **Louise Falzon**: Resources, Data curation, Investigation, Validation, Formal analysis, Writing- Reviewing and Editing. **Yuliang Weng**: Software, Data Curation, Writing- Reviewing and Editing.

## References

1. Uttley L, Quintana DS, Montgomery P, Carroll C, Page MJ, Falzon L, et al. The problems with systematic reviews: a living systematic review. Journal of Clinical Epidemiology. 2023 Apr;156:30–41. doi:10.1016/j.jclinepi.2023.01.011

2. Uttley L, Weng Y, Falzon L. Yet another problem with systematic reviews: a living review update. Journal of Clinical Epidemiology. 2025 Jan;177:111608. doi:10.1016/j.jclinepi.2024.111608

3. Love M, Staggs J, Walters C, Wayant C, Torgerson T, Hartwell M, et al. An analysis of the evidence underpinning the national comprehensive cancer network practice guidelines. Crit Rev Oncol Hematol. 2021/11/29 edn. 2022 Jan;169:103549. Located at: 34838981. doi:10.1016/j.critrevonc.2021.103549

4. Khalid M, Sutterfield B, Minley K, Ottwell R, Abercrombie M, Heath C, et al. The Reporting and Methodological Quality of Systematic Reviews Underpinning Clinical Practice Guidelines Focused on the Management of Cutaneous Melanoma: Cross-Sectional Analysis. JMIR Dermatol. 2023/12/07 edn. 2023 Dec 7;6:e43821. Located at: 38060306. doi:10.2196/43821

5. Streck S, McIntire R, Canale L, Michael Anderson J, Hartwell M, Torgerson T, et al. An Evaluation of Evidence Underpinning Management Recommendations in Tobacco Use Disorder Clinical Practice Guidelines. Nicotine Tob Res. 2022/01/14 edn. 2022 Apr 28;24(6):847–54. Located at: 35023556. doi:10.1093/ntr/ntac012

6. Tanner D, Minley K, Snider K, Hartwell M, Torgerson T, Ottwell R, et al. Alcohol use disorder: An analysis of the evidence underpinning clinical practice guidelines. Drug Alcohol Depend. 2022/01/23 edn. 2022 Mar 1;232:109287. Located at: 35063840. doi:10.1016/j.drugalcdep.2022.109287

7. Bodnaruc AM, Khan H, Shaver N, Bennett A, Wong YL, Gracey C, et al. Reliability and reproducibility of systematic reviews informing the 2020-2025 Dietary Guidelines for Americans: a pilot study. American Journal of Clinical Nutrition. 2025 Jan;121(1):111–24. Located at: Embase; 2036700629. doi:10.1016/j.ajcnut.2024.10.013

8. Murigu A, Wong KHF, Mercer RT, Hinchliffe RJ, Twine CP. Reporting and Methodological Quality of Systematic Reviews Underpinning Clinical Practice Guidelines for Vascular Surgery: A Systematic Review. Eur J Vasc Endovasc Surg. 2025 Nov 13;69(5):674–82. Located at: MEDLINE; 39547389. doi:10.1016/j.ejvs.2024.11.010

9. Moore A, Karadag P, Fisher E, Crombez G, Straube S, Eccleston C. Narrative bias ('spin') is common in randomised trials and systematic reviews of cannabinoids for pain. Pain. 2024/01/16 edn. 2024 Jun 1;165(6):1380–90. Located at: 38227560. doi:10.1097/j.pain.0000000000003140

10. Shyagali TR, Rathore A, Kapoor S, Gupta A, Tiwari A, Patidar R. Evaluation of exaggerated claims in the abstracts of systematic reviews reporting accelerated orthodontic tooth movement: a meta research analysis. Australasian Orthodontic Journal. 2022;38(2):268–77.

11. Steegmans PAJ, Di Girolamo N, Meursinge Reynders RA. Spin on adverse effects in abstracts of systematic reviews of orthodontic interventions: a cross-sectional study (part 2). Syst Rev. 2023/06/21 edn. 2023 Jun 20;12(1):99. Located at: 37340504. doi:10.1186/s13643-023-02269-3

12. Cosgrove L, Mintzes B, Bursztajn HJ, D'Ambrozio G, Shaughnessy AF. Industry effects on evidence: a case study of long-acting injectable antipsychotics. Account Res. 2022/06/01 edn. 2024 Dec;31(1):2–13. Located at: 35634753. doi:10.1080/08989621.2022.2082289

13. Snellman A, Carlberg S, Olsson L. Conflict of interest and risk of bias in systematic reviews on methylphenidate for attention-deficit hyperactivity disorder: a cross-sectional study. Syst Rev. 2023/09/27 edn. 2023 Sep 26;12(1):175. Located at: 37752560. doi:10.1186/s13643-023-02342-x

14. Guo L, Xing X, Ren J, Huang X, Miller S, White H, et al. The reporting completeness of observational systematic reviews and meta-analysis in social science could be improved: a cross-sectional survey. J Clin Epidemiol. 2024 Dec;176:111548. Located at: MEDLINE; 39384164. doi:10.1016/j.jclinepi.2024.111548

15. Rotta I, Diniz JA, Fernandez-Llimos F. Assessing methodological quality of systematic reviews with meta-analysis about clinical pharmacy services: A sensitivity analysis of AMSTAR-2. Res Social Adm Pharm. 2025 Feb;21(2):110–5. Located at: MEDLINE; 39643474. doi:10.1016/j.sapharm.2024.11.002

16. Ferri N, Ravizzotti E, Bracci A, Carreras G, Pillastrini P, Di Bari M. The confidence in the results of physiotherapy systematic reviews in the musculoskeletal field is not increasing over time: a meta-epidemiological study using AMSTAR 2 tool. J Clin Epidemiol. 2024/02/26 edn. 2024 May;169:111303. Located at: 38402999. doi:10.1016/j.jclinepi.2024.111303

17. Salameh JP, Moher D, McGrath TA, Frank RA, Sharifabadi AD, Islam N, et al. Assessing Adherence to the PRISMA-DTA Guideline in Diagnostic Test Accuracy Systematic Reviews: A Five-Year Follow-up Analysis. J Appl Lab Med. 2025 Mar;10(2):416–31. Located at: MEDLINE; 39699177. doi:10.1093/jalm/jfae117

18. Axon E, Kanellopoulou A, Tsokani S, Livingstone N, Hilgart J, Richardson R. Incorporating Non-Randomized Evidence in Cochrane Intervention Reviews: A Scoping Review. J Clin Epidemiol. 2025 May;183:111814. Located at: MEDLINE; 40324697. doi:10.1016/j.jclinepi.2025.111814

19. Helbach J, Wandscher K, Pieper D, Hoffmann F. Cochrane review abstracts are getting longer, but this has no large impact on the reporting quality. J Evid Based Med. 2023/09/07 edn. 2023 Sep;16(3):294–302. Located at: 37674307. doi:10.1111/jebm.12545

20. Wandscher K, Helbach J, Pieper D, Hoffmann F. 'We used standard Cochrane methods' - observational study on reporting according to PRISMA-A in Cochrane review abstracts between 2016 and 2023. J Clin Epidemiol. 2025 Feb;181:111713. Located at: MEDLINE; 39924127. doi:10.1016/j.jclinepi.2025.111713

21. Czajkowski M, Olsson L. Conflicts of interest and industry funding declared in systematic reviews of interventions for six common diagnoses. Scandinavian Journal of Primary Health Care. 2025. Located at: WoS; WOS:001518375300001. doi:10.1080/02813432.2025.2519660

22. Bachelet VC, Lizana FJ, Andrades CO, Carroza B, Gonzalez LRA, Munita P, et al. Estimates for diagnostic accuracy of rapid antigen tests for SARS-CoV-2 in systematic reviews are consistently similar despite poor methodological rigor: a methodological overview. J Clin Epidemiol. 2024 Dec;176:111547. Located at: MEDLINE; 39357662. doi:10.1016/j.jclinepi.2024.111547

23. Siemens W, Bantle G, Ebner C, Blumle A, Becker G, Schwarzer G, et al. Evaluation of 'implications for research' statements in systematic reviews of interventions in advanced cancer patients - a meta-research study. BMC Med Res Methodol. 2023/12/21 edn. 2023 Dec 20;23(1):302. Located at: 38124124. doi:10.1186/s12874-023-02124-y

24. Siemens W, Bantle G, Mahler S, Nothacker J, Stadelmaier J, Bitzer EM, et al. Clinical and methodological implications for research elements in systematic reviews on COVID-19 treatment were often unstructured and under-reported: a metaresearch study. J Clin Epidemiol. 2023/12/11 edn. 2024 Feb;166:111236. Located at: 38072174. doi:10.1016/j.jclinepi.2023.111236

25. Sofi-Mahmudi A, Stojanova J, Vounzoulaki E, Tomlinson E, Pizarro AB, Khademioore S, et al. Cochrane reviews' authorship has become more gender-diverse but remains geographically concentrated: a meta-research study. J Clin Epidemiol. 2025 May;181:13. Located at: WoS; WOS:001447194500001. doi:10.1016/j.jclinepi.2025.111719

26. van Zuilen MH, Saleh MA, Jacobs S, Ferrari CB, Samarakoon T, Duthely LM. A 25-Year Gender Trend Analysis of Authorship in Cochrane Library's Database of Systematic Reviews. Journal of General Internal Medicine. 2025 Apr;40(5):1215–6. Located at: Embase; 2031942128. doi:10.1007/s11606-024-09140-8

27. Wykes T, Garg S, Stahl D, Kostem A, Wilson-Lemoine E. Representation of women in scientific subjects: overview of systematic reviews investigating career progress in academic

publishing with a focus on mental health. BJPsych Open. 2025 Mar 12;11(2):e49. Located at: MEDLINE; 40071959. doi:10.1192/bjo.2024.820

28. Levy MS, Samarakoon T, Smith C, Goo I, Duthely LM, Van Zuilen M, et al. A 20-year analysis of gender trends in oncology authorship in the Cochrane database of systematic reviews. Front Oncol. 2024 Oct 15;14:1450475. Located at: WoS; WOS:001344411500001. doi:10.3389/fonc.2024.1450475

29. Bhat V, Ozair A, Bellur S, Subhash NR, Kumar A, Majmundar M, et al. Inequities in Country-and Gender-Based Authorship Representation in Cardiology-Related Cochrane Reviews. JACC-Adv. 2022 Dec;1(5):3. Located at: WoS; WOS:001492892500012. doi:10.1016/j.jacadv.2022.100140

30. Bondok M, Dewidar O, Al-Ani A, Selvakumar R, Ing E, Ramke J, et al. Inequities in glaucoma research: an analysis of Cochrane systematic reviews and randomized trials. J Clin Epidemiol. 2025 Feb;181:111717. Located at: MEDLINE; 39929324. doi:10.1016/j.jclinepi.2025.111717

31. Walker N, Peden AE, Bestman A, Baffsky R, Willoughby M, Ma T, et al. Interventions that address interpersonal violence experienced by adolescents globally: a systematic review of reviews. Inj Prev. 2024;30(5):363–72. Located at: MEDLINE; 38991717. doi:10.1136/ip-2023-045090

32. Nevill CR, Sharifan A, O'Mahony A, Tahir H, Robinson W, Modha U, et al. Capturing the influx of living systematic reviews: a systematic methodological survey. Journal of clinical epidemiology. 2025 Jul 23;111904. Located at: WoS; MEDLINE:40712752. doi:10.1016/j.jclinepi.2025.111904

33. Young S, MacDonald H, Louden D, Ellis UM, Premji Z, Rogers M, et al. Searching and reporting in Campbell Collaboration systematic reviews: A systematic assessment of current methods. Campbell syst. 2024 Sep;20(3):e1432. Located at: MEDLINE; 39176233. doi:10.1002/cl2.1432

34. Morone G, De Angelis L, Martino Cinnera A, Carbonetti R, Bisirri A, Ciancarelli I, et al. Artificial intelligence in clinical medicine: a state-of-the-art overview of systematic reviews with methodological recommendations for improved reporting. Front Digit Health. 2025 Mar 5;7. doi:10.3389/fdgth.2025.1550731

35. Polizzi A, Leonardi R. Automatic cephalometric landmark identification with artificial intelligence: An umbrella review of systematic reviews. J Dent. 2024;146:105056. Located at: MEDLINE; 38729291. doi:10.1016/j.jdent.2024.105056

36. Reeves S, Patel K, Mukeshkumar K, Naini FB. Cochrane systematic reviews in orthodontics: trends across updates. Eur J Orthod. 2024 Oct;46(5):cjae037. Located at: MEDLINE; 39140148. doi:10.1093/ejo/cjae037

37. Thomas J, McDonald S, Noel-Storr A, Shemilt I, Elliott J, Mavergames C, et al. Machine learning reduced workload with minimal risk of missing studies: development and evaluation of a randomized controlled trial classifier for Cochrane Reviews. Journal of Clinical Epidemiology. 2021 May 1;133:140–51. doi:10.1016/j.jclinepi.2020.11.003 PubMed PMID: 33171275.

38. Uttley L, Matandika L, Falzon L, Tricco A, Labbé C, Byrne JA, et al. Research Culture Influences in Health and Biomedical Research: Rapid Scoping Review and Content Analysis [Internet]. 2024 Apr 3 [cited 2024 Apr 3]. Available from: https://osf.io/mgz93/

39. Uttley L. Research Culture's Role in Contributing to Research Waste: Lessons from Systematic Reviewlution. Exchanges: The Interdisciplinary Research Journal. 2024 Aug 8;11(3):3. doi:10.31273/eirj.v11i3.1539

40. Thomas J, Flemyng E, Noel-Storr A. Responsible use of AI in evidence SynthEsis (RAISE): recommendations and guidance [Internet]. 2024 Aug 29 [cited 2026 May 11]. Available from: https://osf.io/fwaud/overview

41. Uttley L. Review and Analysis of UK Institutional Guidance for Artificial Intelligence in Research. [Internet]. Report commissioned by the UK Committee on Research Integrity and UKRI Research Integrity Working Subgroup.; 2025 [cited 2026 May 11]. Analysis of UK institutional policy and guidance on AI in research – UKCORI. Available from: https://ukcori.org/our-work/artificial-intelligence-and-research-integrity-enabling-trust-and-innovation/analysis-of-uk-institutional-policy-and-guidance-on-ai-in-research/