AIMOS conference abstracts 2024

Information correct as at: 02 October, 2024. Abstracts shown in date/time order.

Tuesday 19 November

Lightning talks 1 - Replicability and reproducibility

Name(s): Cesar Ariza (In person)

Title: A metric-based approach to transparency and reproducibility in Economic research

Abstract: Scientists across disciplines acknowledge the existence of a reproducibility crisis. Contributing factors include selective reporting and limited access to essential research components such as raw data, methods, and codes, which lead to irreproducible research. In response, many journals have implemented transparency policies that force researchers to publish raw data and code and explain the necessary steps to obtain all the published results using these elements. This study aims to develop a metric to evaluate transparency and reproducibility in the context of the data availability policy established by The Review of Economics and Statistics in 2010. These metrics are used to analyse the contents of all the reproducibility material delivered by researchers in the period 2010-2019. Likewise, key features of the research teams (their size and heterogeneity of both countries of affiliation and research seniority) and the papers themselves (publication dates and paper's subject) are examined to evaluate their influence on different levels of transparency. The findings will contribute to assessing how researchers understand transparency and how the interpretation of the data availability policy evolves over time and across different economic subfields.

Name(s): Justine C. Blackwell, Alex O. Holcombe (In person)

Title: How transparent and reproducible are studies that use animal models of opioid addiction?

Abstract: The reproducibility crisis in psychology has cased multiple fields to reckon with the reliability of their own findings. Many of the unfortunate aspects of research design that undermine reproducibility also threaten translation potential (Fergusson et al., 2019). In preclinical addiction research, rates of translation have been described as "disappointing" (Venniro et al., 2020). We examined indices of transparency and accurate and thorough reporting in animal models of opioid addiction from 2019 to 2023. By examining the prevalence of these practices, we aimed to understand whether efforts to improve reproducibility are relevant to this field. We measured the prevalence of transparency measures such as preregistration, registered reports, open data, and open code as well as compliance to the Animal Research: Reporting of In Vivo Experiments (ARRIVE) guidelines. We also measured rates of bias minimisation practices (randomisation, masking, and data exclusion), sample size calculations and multiple corrections adjustments. Lastly, we estimated the accuracy of test statistic reporting using statcheck. The coding of 255 articles revealed poor uptake of all transparency measures and the ARRIVE guidelines. Similarly, levels of bias minimisation practices and sample size calculations were unsatisfactory. In contrast, adjustments for multiple comparisons were implemented in most articles (76.5%). Lastly, half of articles contained non-decision errors and 11% contained decision errors. Our study is the first of its kind in this field.

Name(s): Tatiana Chakravorti (Virtual)

Title: Integrating measures of replicability into scholarly search: Challenges and opportunities

Abstract: Challenges to reproducibility and replicability have gained widespread attention, driven by large replication projects with lukewarm success rates. A nascent work has emerged developing algorithms to estimate the replicability of published findings. The current study explores ways in which AI-enabled signals of

confidence in research might be integrated into the literature search. We interview 17 PhD researchers about their current processes for literature search and ask them to provide feedback on a replicability estimation tool. Our findings suggest that participants tend to confuse replicability with generalizability and related concepts. Information about replicability can support researchers throughout the research design processes. However, the use of AI estimation is debatable due to the lack of explainability and transparency. The ethical implications of AI-enabled confidence assessment must be further studied before such tools can be widely accepted. We discuss implications for the design of technological tools to support scholarly activities and advance replicability.

Name(s): Andrew Vonasch (In person)

Title: Interpreting studies where the main finding replicates but not the control condition

Abstract: We replicated three studies from Hsee (1998). Hsee found that when evaluating choices jointly, people compare and judge the option higher on desirable attributes as better ("more is better"). However, when people evaluate options separately, they rely on contextual cues and reference points, sometimes resulting in evaluating the option with less as being better ("less is better"). We found support for the (surprising) "less is better" effect across all studies (N = 403; Study 1 original d = 0.70 [0.24,1.15], replication d = 0.99 [0.72,1.26]; Study 2 original d = 0.74 [0.12,1.35], replication d = 0.32 [0.07,0.56]; Study 4 original d = 0.97 [0.43,1.50], replication d = 0.76 [0.50,1.02]), with weaker support for the (obvious) "more is better" (Study 2 original d = 0.92 [0.42,1.40], replication dz = 0.33 [.23,.43]; Study 4 original d = 0.37 [0.02,0.72], replication dz = 0.09 [-0.05,0.23]). I will discuss how to interpret the meaning of a study when the "surprising" part replicates but the "obvious" part doesn't. The original effect replicates, but perhaps not because of the mechanism the original studies were designed to test.

Name(s): Anthony Manyara (Virtual)

Title: Minimum sample size in multistakeholder Delphi surveys: replicability of results and recommendations

Abstract: Background The minimum sample size for multistakeholder Delphi surveys remains understudied. Drawing from three large international multistakeholder Delphi surveys, this study aimed to investigate the effect of increasing sample size on replicability of results Methods Resampling with replacement was used to draw random subsamples from the participant data set in each of the three surveys. For each subsample, the median value of all rated survey items was calculated and compared to the medians from the full participant dataset. The median number (and interquartile range) of medians replicated was used to calculate the percentage replicability (and variability). High replicability was defined as $\geq$80% and moderate as 60% and <80% Results The average median replicability (variability) as a percentage of total number of items rated from the three datasets was 81% (10%) at a sample size of 60. In one of the datasets, a $\geq$80% replicability was reached at a sample size of 80. For subgroup analysis based on participant characteristics (e.g. gender, age, professional role), using resampled samples of 20-100 showed that a sample size of 20-30 resulted to moderate replicability levels of 64-77%. Conclusion We found that a minimum sample size of 60-80 participants in multistakeholder Delphi surveys provide a high level of replicability in the results. For Delphi studies limited to individual stakeholder groups (such as researchers, clinicians, patients), a sample size of 20-30 per group may be sufficient.

Name(s): Coralie Williams (In person)

Title: Transparent reporting items for simulation studies evaluating statistical methods in ecology and evolution

Abstract: Simulation studies are important tools for assessing statistical methods by generating data from known processes. However, incomplete reporting can hinder their interpretability and reproducibility, leading to the misuse of statistical methods. While Morris et al. (2019, Stat Med, 38, p2074) recently provided guidance on planning and conducting simulation studies, there are currently no comprehensive reporting guidelines in ecology and evolutionary biology. We propose eleven reporting items for statistical simulation studies, extending on Morris et al. guidance, covering planning, coding, and analysis stages. To illustrate current practices, we surveyed 100 articles in ecology and evolution journals featuring simulation studies. Our findings reveal a need for more transparent reporting to ensure clear evaluation of statistical methods.

Beyond these reporting items, we discuss the benefits of open science tools for enhancing reproducibility. We advocate for the registration of simulation studies to improve planning, reporting, and collaboration. Our goal is to initiate discussions on improving the reporting standards for simulation studies, enhancing reproducibility, and encouraging as a side effect better design and conduct in statistical method research.

Lightning talks 2 - Reviews and meta-analysis

Name(s): Joshua Wang, Hayley Moody (Virtual)

Title: Examining Language Inclusion Intentions in Scoping Reviews

Abstract: Research published in languages other than English (LOEs) is often ignored in evidence syntheses, marginalising diverse knowledge. While the extent of LOE inclusion, and the associated attitudes of LOE inclusion amongst authors of systematic reviews has been well characterised, LOE inclusion in other evidence synthesis forms has yet to be explored. Scoping reviews, in comparison to systematic reviews, examine a broader range of sources to build a conceptual summary of a field of inquiry, making LOE literature an important source of information for scoping review authors. To understand the current state of LOE inclusion intentions in scoping reviews, peer-reviewed, PubMed-indexed scoping review protocols published in from 01-Jan-2024-11-Aug-2024 were analysed for LOE inclusion. Overall, LOE inclusion intentions and attitudes were diverse, with just under half of the 249 protocols analysed including an LOE. Many LOE-inclusive articles relied on the authorship team's own LOE proficiency to gather evidence. Machine translation was also intended to be used in one quarter of the LOE-inclusive protocols. Approximately 70% of LOE-exclusive protocols did not state that they would exclude LOEs at the screening stage in order to report the number of LOE articles excluded. Together, this analysis demonstrates the need for increased LOE inclusion and reporting guidelines for scoping reviews, as well as the importance of analysing LOE inclusion in other forms of evidence synthesis.

Name(s): Matthew Page (In person)

Title: Investigation of the replicability of systematic reviews of the effects of interventions: the REPRISE study

Abstract: Objectives: To evaluate the extent of variation in results when we independently replicate original systematic reviews. Methods: We randomly selected 10 systematic reviews of the effects of interventions indexed in November 2020, in which methods and results of the searches and meta-analyses were reported completely. Two librarians re-ran the original search strategies (limited to the original time frame) and recorded the number of records retrieved. Two investigators independently collected from each study report the data needed to replicate the first meta-analysis and calculated a summary effect estimate and its 95% confidence interval. We calculated for each database search how often the difference between the original and replicated number of records yielded was >10%. We compared the original and recalculated meta-analysis results and judged whether the observed difference was meaningful (i.e. change in statistical significance or direction). Results: Across the 10 systematic reviews, 44 bibliographic database searches were re-run. The difference between the original and replicated number of records yielded was > 10% for 26 (59%) databases. Meta-analysis results were not fully replicable in four (40%) cases, although the difference was not considered meaningful in any case. Conclusions: The findings of original systematic reviews are not always consistent when the methods are repeated. Greater clarity and detail about the search methods, study eligibility criteria and methods for selecting study results may enhance replicability of systematic reviews.

Name(s): David Borg (co-authors: Sarah R Kane and Anna Lene Seidler) (In person)

Title: Meta-analysis Studies in Sports Science and Medicine

Abstract: Background: In 2023, 1,800 meta-analysis studies were published in the field of sport. Considering the methodological and statistical expertise required to conduct meta-analysis, and in light of the limited number of researchers with such expertise in the field, the high volume of published meta-analysis raises potential concerns of their quality, and hence of possible research waste. In this study, we explored the current state of meta-analysis studies in Sports Science and Sports Medicine.

Methods: We included all 2024 meta-analysis studies, published in six popular, quartile one ranked Sports Science and Sports Medicine journals. We extracted data on: registration, reporting, software, whether certainty of evidence was considered, risk of bias, and data/code sharing.

Results: Preliminary analysis of 30/87 meta-analysis studies showed that 27 studies were prospectively registered. R software was most common (n=16), followed by RevMan (n=7), CMA (n=3), and Stata (n=3). Thirteen studies rated the certainty of evidence, and 26 studies assessed the risk of bias. Four studies considered the results according to levels of the quality appraisal or risk of bias outcomes. Five studies shared their data and code.

Conclusion: Our study suggests that a large proportion of sport researchers still use analytically inflexible software that may prevent the use of models needed to account for more complex data structures (i.e., multiple effect sizes from the same study). There is a lack of consideration of study quality when doing meta-analyses, and data and code sharing remains low.

Plenary - Where next for metascience? Prospects, pitfalls and possibilities for the next five years

Name: James Wilsdon (Chair: Fiona Fidler)

Title: Where next for metascience? Prospects, pitfalls and possibilities for the next five years

Abstract: After a decade of steady growth, metascience is exploding into the mainstream. Its simple premise – that we should turn the methods and tools of research back towards analysing and improving the research system itself – is now increasingly accepted by governments, funding agencies and research communities themselves. Of course, these agendas aren't new: metascience builds on rich interdisciplinary currents of research into R&D systems. But there is also a growing cadre of researchers, policymakers and practitioners – in universities, R&D-intensive companies, funding agencies, private labs and foundations – who are now deploying advanced methods and data to investigate and improve research systems, cultures and decision-making. Worldwide, we see new initiatives, investments and alliances being set up to strengthen the field. To give just six examples: - UK, where the government has established an in-house Metascience Unit and UK Research and Innovation (UKRI) has set up a metascience grants scheme. - Canada, which is launching a new cross-agency funding call for meta-research; - Germany, where the Volkswagen Foundation has a programme for "Researching Research"; - USA, where the NSF is complementing its "Science of Science" funding program by working with others to experiment with new funding methods; - China, where the recent Shanghai Declaration for the Science of Science called for more sustained and ambitious investment in these fields; - Ireland, where there is a concerted drive to embed metascience in the new combined funding agency, Research Ireland. In this talk, drawing on a new study by RoRI and CWTS-Leiden, James Wilsdon will take us on a whistle-stop tour through this changing landscape, and highlight a few metascientific movers and shakers, hotspots and lacunae. He will argue that metascience is less a discipline, and more an orientation or mode of engaging with questions that almost all researchers encounter at some point in the networks and institutions they inhabit. Finally, he will offer a few thoughts on the metascientific possibilities of the present moment, and outline five priorities for the next five years.

Parallel sessions

Category: Discussion group

Name(s): Ze Freeman (In person)

Title: Developing transparency in clinical psychology research

Abstract: Clinical psychology research is driven by a need to establish effective therapies and mechanistic understandings of mental illness to improve patient outcomes. Despite significant funding in the area over many decades, the evidence base for most therapies are weak and outcomes for people experiencing difficulties, particularly those with severe mental illnesses, remain poor. Incentives for clinician-researchers may appear to be more aligned with principles of transparency and openness than non-clinicians, but this does not seem to be the case when looking at the lack of openness in the literature. This session will focus on the life-cycle of

clinical research, from theoretical underpinnings to dissemination of findings. Areas for metascience to focus on, including appraising clinical trial reporting, assessing spin, and evaluating theory use will be discussed.

Category: Discussion group

Name(s): Soumyadeep Bhaumik (In person)

Title: Research priority setting as a pro-justice tool in the knowledge ecosystem of medicine and health

Abstract: Traditionally, the success of research is measured by its value within the research community, with limited attention to its broader social impact. This is changing and research priority setting (RPS) is a tool being used to address research waste. However, RPS currently focus on only what research questions it addresses. In the discussion workshop we seek to explore how RPS exercises can address the problem of unequal representation, concentration of power, epistemic injustice, by modulating and nudging the knowledge ecosystem towards fairness and justice

Category: Discussion group

Name(s): Adrian Barnett (In person)

Title: What should AIMOS do next?

Abstract: AIMOS was founded in June 2019. In our first five years, we have been an active scholarly society and have advanced the field of meta-research. This discussion group aims to gather your thoughts and ideas for what AIMOS should do in the next two years. Some potential goals are to: Improve our finances and sustainability, Increase membership, Increase our profile and influence, Run an annual conference. Are there other important goals? What actions should we take to achieve these goals? Should we pursue deductible gift recipient status? Should we raise our profile in the global south? How can we best balance these ideas against our limited resources? This discussion forum will be a great opportunity for our community to shape the future of AIMOS. I will start with a brief outline of our previous activities, member numbers and spread, and financial status. Please share your ideas for how AIMOS can continue to be an active society that improves the quality of research.

Parallel session

Category: Discussion group

Name(s): Patrick S. Forscher (Virtual)

Title: Developmental meta-research: The intersection between meta-research and global development

Abstract: Over the past three years, the CREME Developmental Meta-Research team has been investigating intersections between meta-research and a seemingly very different field: global development. Housed within Busara, a non-profit dedicated to using behavioral science to achieve development goals, the team has both consulted internal and external partners and attempted to use meta-research methods to ongoing projects within the organization. We have discovered that many communities in development are already interested in understanding and improving the research process. Some of these are concerned with the current central focus of meta-research — research credibility – but some are concerned with improving other aspects of research, such as the way participants are treated, the evidence-to-policy pipeline, power differences across the Global North and Global South divide, and research capacity. In this session, we will discuss the results of our investigations, showcasing what we argue can be a fruitful and productive subdiscipline of meta-research that can exist at the intersection of meta-research and global development. We call this subdiscipline developmental meta-research.

Category: Discussion group

Name(s): Tim Fellows (Virtual)

Title: Octopus: Examining how a new publishing model could improve research culture

Abstract: Octopus.ac (www.octopus.ac) is a new open research publishing model, designed to encourage, enable, and reward best practice. The platform is open source and entirely free to read and publish on.

We believe that many of the problem in academic publishing stem from one principal issue: that journals are being pulled in two different directions – the dissemination of findings to practitioners and general audiences, and being the primary research record of what has been done, when and by whom, in detail, for the benefit of specialists. This leads to key content being dismissed to supporting appendixes, while researchers try to write their results in a highly-narrative, attention-grabbing way which maximises 'impact'.

Octopus seeks to sit alongside journal articles by using smaller publication units which more closely align with the research process to fulfil the role of the primary research record for the research community. The platform is designed to create a new culture of collaboration and recognition, improving access to research outputs and resetting the academic incentive structure to reward best practice and recognise specialisation.

In this discussion we will examine the opportunities and challenges Octopus faces in bringing about positive change in the research culture, informed by both the experiences and opinions of attendees, and the outcomes of a report published last year that examined perceptions in research publishing: https://www.octopus.ac/research-culture-report

Category: Hackathon

Name(s): Max Korbmacher (Virtual)

Title: Translation and extension of the glossary of open science terms

Abstract: We propose an AIMOS 2024 hackathon focused on expanding and translating the Framework for Open and Reproducible Research Teaching (FORRT) Open Scholarships Glossary (https://forrt.org/glossary/english/), a valuable resource aimed to create a centralized, open-access collection of open scholarship terminology.

This hackathon directly contributes to the Open Scholarship movement by improving the accessibility and comprehensiveness of the FORRT glossary. The project welcomes researchers from all disciplines to ensure the glossary reflects the varied and current understanding of Open Scholarship terminology within the research community. Moreover, the translations will broaden the glossary's reach, making it a valuable tool for researchers worldwide.

Participants will directly work on the Glossary, either in the translation or the extension team. All contributions will be credited by being mentioned on the website and by an invitation to participate in resulting publications. We will also encourage hackathon participants to continue working on the project after the hackathon is over.

Mini-notes - Research training

Chair: Chair Mark Hooper

Title: Research training

Abstract: Konrad Kording (C4R): Introducing the Community for Rigor: A Free Rigor Education Network. This session will demonstrate educational tools developed by the Community for Rigor (C4R), an NINDS-funded initiative that makes a free, online network of resources providing instruction in the principles of rigorous research. Participants will hear remarks from Principal Investigator Konrad Kording and discuss the importance of conducting awesome and reliable scientific research. Then, participants will briefly participate in a group-based interactive demo to engage with C4R's approach to rigor education. Natalie Evans (The Embassy of Good Science): The Embassy of Good Science: Training 'Good' Researchers. How can we make researchers care about research integrity? How can we put the learner at the center of our teaching? How can we make an impact at scale? This presentation will describe the approach to training researchers taken by The Embassy of Good Science, a European funded initiative for making research integrity and ethics accessible and engaging. We will introduce the 'VIRT2UE' programme, a training approach that focuses on researchers' virtues, which – with 600 trainers across Europe – has had a large scale uptake. We will also describe the possibilities for educators to curate and develop their own engaging e-learning materials tailored to the needs of their learners on The Embassy of Good Science wiki platform. Flavio Azevedo (Framework for Open and Reproducible Research Training): FORRT: Lessons and Resources from an international and educationally-oriented open science organisation devoted to DEI and Social Justice aspects.

Plenary - Structural Blind Spots: Why We Know So Little about Wrongful Convictions

Name: Barbara O'Brien (Chair: Justice Belinda Baker)

Title: Structural Blind Spots: Why We Know So Little about Wrongful Convictions

Abstract: Barbara O'Brien will draw on her legal practice and research, including editing the U.S. National Registry of Exonerations, to discuss why we know so little about wrongful convictions. She will argue that many reasons are due to the same sorts of mechanisms and institutional interests that undermine transparency across scientific disciplines. These include a lack of uniformity across jurisdiction in keeping very basic records, various doctrines that give government actors immunity from lawsuits, asymmetry in evidentiary discovery between civil and criminal cases, and judicial commitment to "finality" in convictions.

Wednesday 20 November

Mini-notes - Some good news from meta-research

Name: Chair Alex Holcombe

Title: Some good news from meta-research

Abstract: David Mellor (Open Science Framework): Good news everyone! How meta-research has improved policy and reform efforts. Meta-scientific research has played an instrumental role in both improving the policies and practices across multiple disciplines, but has also served as an important tool to measure progress and to correct reform efforts. In the former, large scale replication efforts and other efforts to use existing scholarly literature have provided empirical evidence to the theoretical problems that had been discussed for decades. In the latter, policies that encourage open science practices such as data sharing, study registration, and Registered Reports have identified shortcomings that result in tangible improvements to how these policies are implemented or enforced. This presentation will emphasize the connection between meta-scientific studies and changes that have occurred because of them. Tammy Hoffmann (Bond University): TIDIER. Ivan Oransky (Retraction Watch): Anita Bandrowski (University of California): SciScore.

Plenary - Metascience insights into the dynamics of science communication: the case of STEM cell research

Name: Joan Leach (Chair: Rose O'Dea)

Title: Metascience insights into the dynamics of science communication: the case of stem cell research

Abstract: Concepts of openness and reliability have been developed by metascience scholars and applied across fields to bring out epistemic as well as social vulnerabilities in the practice of science. Calls to make science more open and reliable gesture towards contemporary efforts to, respectively, both expand how science is undertaken whilst also bolstering its broader legitimacy. This paper will look at the recent history of stem cell research and, drawing from empirical data from the EOAR project (Enabling Openness in Australian Stem Cell Research), will consider how openness and reliability are understood, positioned and valued by researchers in this field. Further, this talk will review how these concepts impact the public communication of this area of science and ask how they should be imagined and enacted going forward. The arguments of the paper will be developed from stem cell science, but will hopefully raise issues of a much more general nature, especially about the public communication of metascience issues.

Lightning talks 3 - Policy and practice

Name(s): Nofar Sheffi (In person)

Title: Challenging (Con)Census

Abstract: Recent debates about the upcoming census have highlighted persistent questions about inclusion, representation, and capture, including in relation to statistical methods and their employment in census-taking. What questions should be included? How should these questions be framed and structured? What response options should be provided? What labels should be used? How should the collected data be coded? My proposed Lightening Talk questions the assumptions underpinning and grounding the census as an enterprise. For example, the Standard for Sex, Gender, Variations of Sex Characteristics and Sexual Orientation Variables (2021) explains that the collected information can be used by government and 'academic and private sector organisations' for a range of purposes, including: informed decision making and planning; policy formulation and monitoring; social, population and economic research and analysis; [and] program provision and evaluation (e.g. health services)'. My talk will draw on political economy and critical investigations of data-based governance and raise a provocative set of questions about the census, its design, and its use, reflecting on the prevalent utilisation of census data by corporations that design, employ, and supply machine learning (ML) and artificial intelligence (AI) tools.

Name(s): Zachary Kendal (In person)

Title: The impact of Read & Publish agreements on Australia's open access landscape

Abstract: After some pilots with smaller publishers in previous years, 2022 saw Australia's first open access publishing agreements with major scholarly publishers. Now, in 2024, there are 25 agreements on offer through the Council of Australian University Librarians (CAUL), which negotiates and manages the agreements on behalf of Australian universities. Most of these are Read and Publish (R&P) agreements built on existing library subscriptions, and there are now agreements in place with the 'big five' commercial publishers, as well as several university presses and society publishers. These CAUL R&P agreements covered about 430 papers in 2021, over 11,000 in 2022, and almost 23,000 in 2023. However, few of the major agreements cover publishing in fully open access (gold) journals, resulting in an explosion in 'hybrid' open access publishing in subscription journals in recent years. This lightning talk will examine the impact of the CAUL R&P agreements on the open access landscape in Australia, including authors' uptake of different open access pathways and compliance with funder mandates. It will also reflect on the challenges and risks associated with the R&P model, issues around integrity and publishing ethics, and the need to foster diverse approaches to open scholarship.

Name(s): Cooper Smout (In person)

Title: WISDOM: An open science framework for community-driven value accounting

Abstract: In this lightning talk, I'll introduce WISDOM (Weighted Information Schema for Distributed Open Merit) and share data from our pilot at the AIMOS 2023 conference. I'll begin with an overview of the model, showing how it can adapt to diverse use cases and facilitate the collective valuation of contributions according to community-selected values. I'll highlight the key innovation: the use of pairwise comparisons as both an inclusive review protocol and a 'standard unit of contribution', which serve to peg the value(s) of all other contributions in a decentralised and autonomous manner. I'll share data from our recent prototype, where we used WISDOM to recognise and reward contributions to the AIMOS 2023 conference. I'll show how we empowered the AIMOS community to: (1) Record diverse contributions, including many that might otherwise go unnoticed; (2) Review contributions according to community-selected dimensions; (3) Recognise contributions for their diverse qualities; (4) Reward contributors for the value they provided; and, (5) Respect reviewers and contributors for their demonstrated expertise. I'll present our major analyses and system upgrades, showing how this circular process of meta-research and iteration could produce a highly efficient, values-aligned research ecosystem. I'll close with an invitation to nominate AIMOS 2024 contributions online and attend our follow-up hackathon, thus contributing to the latest iteration of the WISDOM experiment.

Name(s): Jennifer Ding, Jack Conner, and William L. D. Krenzer (Virtual)

Title: What Keeps Researchers from Publishing Registered Reports?

Abstract: Registered reports are a recently developed publication method that allows researchers to submit a detailed proposal for their project and receive early feedback and an in-principle acceptance to a journal prior to data collection. By accepting papers prior to data collection, registered reports should reduce publication bias, while promoting transparency and quality research practices. Despite the benefits, however, registered reports have not been widely adopted. To understand why, this study analyzes data collected from 9 focus groups, during which researchers were asked about their knowledge and thoughts on registered reports. After a detailed presentation explaining the process of publishing a registered report, opinions were collected again. A set of coders summarized participant statements, creating a condensed, consistent set of data to complement the raw data. While the former provides thematic information, the latter gives a more nuanced look at opinions. Using R, Coders analyzed each dataset, generating word clouds and performing cluster and sentiment analysis. Coders also examined pre- and post-presentation attitudes to explore how future registered report outreach programs can be improved. Our results demonstrate that the major concern surrounds the timeline of registered reports, while other deterrents include wariness about peer reviewer quality and a general reluctance to stray from tradition. By identifying current opinions and concerns, our data can help direct efforts in expanding the implementation of registered reports.

Parallel sessions

Category: Discussion group

Name(s): Julia Bottesini, Priya Silverstein, Sebastian Karcher (Virtual)

Title: Leveraging the Journal Editors Discussion Interface (JEDI) for Metascience

Abstract: Interested in metascience and scientific publishing? Come pitch a metascience project that takes full advantage of the Journal Editors Discussion Interface (JEDI; dpjedi.org) network! JEDI is a community of nearly 500 social science journal editors and those interested in improving scientific publishing. JEDI aims to facilitate discussion and build a fund of collective knowledge on editorial practices and their intersection with open scholarship. With access to hundreds of social science journal editors and their journals, JEDI is uniquely positioned for conducting metascientific research on scholarly publishing. The JEDI team will present potential project ideas and hear yours, followed by an in-depth discussion of the opportunities and complications of different potential metascientific study methods. Although not required, the JEDI team is open to the potential for this session to evolve from discussion into a hackathon to zoom in on one idea and sketch out a rough study proposal!

Category: Discussion group

Name(s): Kathy Bowrey; Jennifer Byrne; Pranujan Pathmendra (In person)

Title: Policy and practice for post-publication corrections- biomedicine meets the law

Abstract: The capacity to correct the published record is important for evidence-based research and scholarship that can be translated to practice. Nonetheless, it is widely recognised that post-publication corrections are not keeping pace with rising numbers of publications and increasing publication capacity, particularly through the uptake of gold open access publishing models. This discussion group will include biomedical researchers with experience in describing flawed publications to many different journals and publishers, and a legal expert with expertise in publishing agreements and author's rights. The biomedical researchers will describe the low rates of engagement that they have experienced when describing flawed cancer research papers to different journals and publishers over the last 10 years. The legal expert will describe publishing terms and practices that may inhibit post-publication, corrections and assess whether these create significant barriers in practice. The group will then discuss how post-publication correction capacity could be improved, and ways to test suggested approaches in real world settings.

Category: Workshop

Name(s): David Flanders Tuke, Cameron Neylon, Karl Huang, Pablo Franco Ulloa (Virtual)

Title: Building Better University Research Policy Through AI-Enhanced Virtual Experimentation

Abstract: This workshop aims to explore the potential of AI-driven virtual experimentation in shaping more effective university research policy. By bringing together researchers, policymakers, and technologists, we will delve into how AI can be leveraged to optimize research funding, resource allocation, and collaboration. Participants will engage in hands-on activities to develop and test virtual experiments focused on real-world research challenges. This workshop aligns with the principles of meta-research by promoting rigorous evaluation of research practices and fostering open science through knowledge sharing and collaboration. By the end of the workshop, participants will have a deeper understanding of the potential of AI-enhanced experimentation and be equipped to apply these insights to their own research and policy work.

Lightning talks 4 - Integrity and trust

Name(s): Pranujan Pathmendra (presenter), Jennifer Byrne (In person)

Title: Do PubPeer posts contribute to published corrections of flawed cancer research articles?

Abstract: Science is expected to be self-correcting, yet journals and publishers often do not act on third-party requests to correct flawed publications. The PubPeer platform enables post-publication commentary on research articles, where some PubPeer posts may contribute to post-publication corrections. To examine a possible relationship between PubPeer posts and post-publication corrections, we studied responses to PubPeer posts about nucleotide sequence reagent errors in 107 original articles in cancer journals Oncogene (n=21 articles) and Molecular Cancer (n=86 articles). All errors in these articles were described in a bioRxiv

preprint in February 2023, which was also sent to the publisher by email in February 2023. Standardised and signed PubPeer notices were then written to describe reagent errors in each article. All articles were ranked according to PubMed identifiers and assigned to one of 2 groups using alternate PubMed identifiers. Group numbers were decided by coin toss. PubPeer notices for group 1 (n=54 articles) were posted on 5th April 2023 and group 2 notices (n=53 articles) were posted on 4th October 2023. Overall, 28% (15/54) group 1 and 21% (11/53) group 2 articles, respectively, received author responses through PubPeer, mostly during the month PubPeer notices were posted. Two group 1 Molecular Cancer articles received published corrections acknowledging reagent errors 12 or 16 months after PubPeer posts. In summary, whereas some PubPeer posts describing reagent errors resulted in author engagement, few articles were corrected by Molecular Cancer or Oncogene.

Name(s): Cesar Ariza (In person)

Title: Factors influencing retraction time lag

Abstract: The number of scientific paper retractions has surged in recent years, with some sources reporting more than 10,000 papers being retracted in 2023 alone. This alarming trend highlights the urgency of identifying factors that influence the retraction time lag (RTL), that is the period between a paper's original date of publication and its retraction. Previous research has explored RTL, but most studies are limited to specific disciplines or short observation periods. This study aims to overcome this limitation by expanding the scope of RTL analysis across a broader range of scientific disciplines and an extended period. The sample analysed includes more than 7,000 papers from the five scientific domains defined by Science-metrix and published over two decades. Additionally, the analysis goes beyond the retractions' reasons and considers the influence of factors such as research teams' characteristics (such as their sizes and heterogeneity of both countries of affiliation and research seniority) and massive simultaneous retractions on RTL. The findings will contribute by improving the early detection of papers at risk of retraction and enhancing research integrity and accountability across the scientific community.

Name(s): Annie Whamond (In person)

Title: How do psychology journals handle post-publication critique?

Abstract: Post-publication critique, such as letters-to-the-editor, can contribute to the validity and trustworthiness of scientific research. To understand how these critiques are handled in the field of psychology, we conducted a cross-sectional analysis of the policies and practices of post-publication critique at randomly-selected and prominent (high impact factor) psychology journals. We found that, in 2023, submission of post-publication critique was explicitly available at 23% (95% CI [16% to 32%]) of randomly sampled psychology journals (N =100) and at 38% of the most prominent psychology journals (N = 100). Journals sometimes imposed limits on the length and time allowed to submit critiques. Based on manual inspection of random samples of empirical articles published in 2020 (N = 101 articles per sample), we estimate the prevalence of post-publication critique to be 0% (95% CI [0% to 3.7%]) in psychology journals generally and 1% (95% CI [0.2% to 5.3%]) in the most prominent psychology journals. These results indicate that the policy and practice of post-publication critique is seriously neglected in psychology. To enhance accessibility of critiques and improve accountability, we recommend psychology journals explicitly offer options for post-publications critiques without imposing restrictive limitations.

Name(s): Rebekah McWhirter (In person)

Title: Improving the quality of ethical review for trustworthy research

Abstract: Research depends upon sustained public trust, and should ideally be worthy of that trust. Ethical review of human research contributes to the maintenance of trustworthy research through assessing for adherence to a set of minimum ethical standards. However, inconsistency in application of these standards both within and between human research ethics committees (HRECs) raises questions about the quality of ethical review processes, particularly in relation to quasi-legal decisions such as s 95A waivers of consent. This causes substantial frustration and wasted resources, especially for multisite research projects. Here, we report on the progress of a Shared Ethical Decision-making (ShED) project that aims to serve as a benchmarking exercise for Australian HRECs. This will provide an evidence base for improved guidance to

HRECs making waiver of consent decisions by identifying how HRECs are applying the current guidelines and by incorporating consumer perspectives to better align HREC decision-making with public expectations.

Name(s): Lyle Gurrin (In person)

Title: Using ChatGPT-4o to semi-automate trustworthiness assessment of randomised controlled trials: A case study

Abstract: PURPOSE: Randomised controlled trials (RCTs) are the cornerstone of evidence-based medicine. Many published RCTs are, however, not based on real data. This compromises the reliability of systematic reviews and meta-analyses, leading to misinformed clinical guidelines. Methods to detect problematic RCTs are time-consuming and labour-intensive. The use of artificial intelligence large language models (LLM) has the potential to accelerate the data collection and statistical analysis needed to assess the trustworthiness of published RCTs. METHODS: We present a case study using ChatGPT powered by OpenAI's GPT-4o to apply the TRACT checklist to one RCT paper. RESULTS: ChatGPT successfully (1) processed the PDF of the selected publication; (2) responded to specific prompts; and (3) addressed items in TRACT by providing precise 'yes' or 'no' answers. Synthesis of information from the paper and online resources depended on ease of access, providing further impetus for the Open Science initiative. A comparison of results generated by ChatGPT, and the human assessor showed an 84% level of agreement of (16/19) TRACT items. ChatGPT reproduced summary data tables as Microsoft Excel worksheets and reorganize the data, with three out of four extracted tables achieving an accuracy score of 100%. CONCLUSION: ChatGPT demonstrates potential in semi-automating the trustworthiness assessment of RCTs, though in our experience this required repeated prompting from the user. Ultimately, we seek a completely automated process for large volumes of papers.

Mini-notes - AIMOS top-up scholarship winners

Chair: Chair Pranujan Pathmendra

Title: AIMOS top-up scholarship winners

Abstract: Beth Clarke (University of Melbourne): How does science generate knowledge: An empirical investigation of the state of social and personality psychology. Adeyemi Adetula (Université Grenoble Alpes): Increasing the generalizability of psychology by making it less WEIRD. Emma Wilson (The University of Edinburgh): A meta-research framework for evaluating evidence from animal models of neurodevelopmental conditions. Maximilian Primbs (Radboud University): A many-lab many-analyst investigation of the generic masculine effect.

Thursday 21 November

Mini-notes - MetaROR launch

Name: Chair TBC

Title: MetaROR launch

Abstract: Simine Vazire Editor-in-Chief of Psychological Science. Ginny Barbour Editor-in-Chief of Medical Journal of Australia. Sujatha Raman Associate Editor at the Journal of Responsible Innovation and a member of the Editorial Collective of Engaging Science, Technology, and Society.

Parallel sessions

Category: Discussion group

Name(s): Soumyadeep Bhaumik (In person)

Title: Why do funders not fund meta-research in health and medicine?

Abstract: Meta-research (research on research) serves an inherent value, as it aims to improve the process and conduct of research itself. However, funding is typically targeted towards specific diseases, issues or population groups. The discussion aims to understand the underlying reasons behind it and explore the potential to co-develop a position paper/ call to action on the issue targeting funders globally. Participants will jointly do mind mapping, to explore different facets of the issue, branching into related aspects including but not limited to definition, perceived value, uncertainty, awareness, power dynamics, and complexities around meta-research.

Category: Hackathon

Name(s): Fiona Fidler, Ginny Barbour, Fallon Mody (In person)

Title: Developing a repliCATS peer review protocol for medical research

Abstract: For over 5 years, the repliCATS project has been eliciting forecasts of replicability and conducting rapid, collaborative peer review in the social sciences. The project has evaluated over 4000 papers in psychology, sociology, education, economics, criminology, political science, and public administration, and provided peer review training for over 1000 researchers in those disciplines. In 2025, it will turn its attention to evaluating health and medical research, and particular providing peer review training for reviewers at the Medical Journal of Australia (MJA). The first step in this new endeavour involves creating a review protocol that is tailored to medical, rather than social science, research. This hackathon will take our existing repliCATS rubric as a starting point, and work on translating and/or replacing items. The goal is to create a modified repliCATS review protocol, ready to be used in peer review training at MJA next year.

Category: Hackathon

Name(s): Cooper Smout, Anna Finanne, Matthew Ruby, Aidan Tan (tentative), Ginny Barbour (tentative) (In person)

Title: Reviewing, recognising and rewarding diverse contributions to the AIMOS conference

Abstract: The Weighted Information Schema for Distributed Open Merit (WISDOM) is an open science and meta-research framework for estimating the relative value of contributions to diverse communities. It features an inclusive review protocol and a decentralised, autonomous process to estimate contribution value, which can be used to recognise, reward and respect contributors for the value they create.

In this hackathon, we'll use WISDOM to review and recognise diverse contributions to the AIMOS 2024 conference. We'll begin with a demonstration of our prototype for AIMOS 2023, showing how contribution data were processed through each stage of the model (Record, Review, Recognise, Reward, and Respect) and sharing our major learnings from conducting this research.

Next, we'll start work on the AIMOS 2024 prototype. Participants will be invited to nominate contributions and dimensions for rating via an online survey. We'll edit the list of contributions and select a representative sample for review. We'll discuss pros and cons of different rating dimensions, then choose a set of dimensions on which to rate AIMOS 2024 contributions. If our numbers permit, we may split into breakout groups to complete these tasks in parallel.

We'll also aim to complete some reviews during the hackathon itself, so please bring a laptop if you have one available. We'll close the session with a round of reflections on the process, an overview of next steps, and an invitation to contribute further to the project.

Lightning talks 5 - Data collection, sharing and re-use

Name(s): Owen Forbes (In person)

Title: Impact-to-Effort Analysis in Research Prioritisation: Data-driven Tools for 'Collecting Smarter' in Ecology

Abstract: Declining data availability from biological collections is a growing concern, impacting our ability to understand ecological phenomena and respond to environmental crises. Specimen-based data provide critical insights on spatiotemporal trends in ecology. Resource constraints necessitate a 'collect smarter' approach for allocating resources in research operational strategy. This reflects a broader need for data-driven approaches to implement systematic methods for prioritising projects that optimise the impact-to-effort ratio in addressing urgent challenges. We propose a research prioritisation framework for assessing dimensions of research projects including complexity, effort, urgency, and expected impact, building on related ideas in business research, health economics, and adaptive design.

We present a case study for research prioritisation in collections science for Australian specimen-based data from the Global Biodiversity Information Facility (GBIF). We use optimal design tools to build a prototype interactive platform for ecology researchers to apply data-driven collection sampling design. Previous approaches to formal research prioritisation in ecology have used Bayesian decision-theoretic tools from optimal design literature regarding value of information (VoI). We suggest combining spatial VoI metrics with derived indices for spatial distribution of ecological problems, as a measure of 'need for information' (NfI) to complement these perspectives and distinguish between dimensions of ecological problems and research projects.

Name(s): Kylie Hunter (In person)

Title: Introducing a framework for secondary use of clinical trials data in health research

Abstract: Background: Data sharing can add tremendous value to existing clinical trials data. The Health Studies Australian National Data Asset (HeSANDA) program by the ARDC built a national infrastructure platform to allow researchers to access and share data from health studies. Yet, guidance around how these data may be used for secondary research was lacking.

Objective: To address this, ARDC sought to develop a theoretical framework for the use of clinical trials data for secondary research.

Methods: The framework was derived from research papers, consultation with key stakeholders, and a discussion group among the meta-research community at AIMOS 2023 to identify secondary research scenarios.

Results: The resulting framework for secondary data use comprises four different scenarios: 1) Evidence synthesis (including aggregate and individual participant data meta-analysis); 2) Secondary analyses (e.g. descriptive, economic, prognostic, predictive, exploratory analysis); 3) Reproducibility, replication and validation (to verify the accuracy, validity, and trustworthiness of the scientific findings of original studies); and 4) Education and Methods Development (e.g. teaching/learning data analysis methods, demonstrating new statistical methods). Full details, including step-by-step user guides and illustrative case studies are available in a published report (Hunter et al. 2024 http://doi.org/10.5281/zenodo.12768050). We anticipate this

framework and report will maximise the scientific value of the HeSANDA platform and inform secondary research in general.

Name(s): Aidan Tan (In person)

Title: Researcher adherence to data sharing policies of journals

Abstract: Rates of data sharing are low. Data sharing might be improved if data sharing policies of journals were prominent and actionable. Our objective is to systematically assess researcher adherence to data sharing policies of journals.

This is a cross-sectional meta-research study. We included all original research published in 2022 in high-impact clinical medicine journals (the five journals with the highest impact factor for all 59 fields of clinical medicine) with a data sharing policy which either recommended or required data sharing. We manually collected data on the initial data sharing plans in study registration records and published study protocols, and the final data sharing decisions in published original research.

We have included 94 journals, from which we have included a total of 13,387 studies which met the inclusion criteria. Only a minority of studies had a study registration record or published study protocols (12.3%, 1640/13387). Of those which did, more than half stated their data sharing intentions in their publicly available research protocol (58.2%, 954/1640), and of these, less than a third intended to share data (31.4%, 300/13387). More than half of studies stated their data sharing decisions in their publication (58.2%, 7794/13387). Of those that did, most intended to share data (87.1%, 6790/13387). Those that did not intend to share data (12.9%, 1003/13387) most commonly did so for no provided reason (26.9%, n=270), because they were not the sole data custodian (20.2%, n=203) or because they planned to publish only summary findings (19.7%, n=198).

Name(s): Jannik Aagerup (In person)

Title: State of Play in IPD Part 1

Abstract: State of Play in individual participant data meta-analysis (IPD-MA) Background. Systematic reviews and meta-analysis are commonly conducted using aggregate data. Performing a meta-analysis on individual participant data (IPD-MA) enables more powerful and detailed analysis with the potential for greater impact on policy and practice. IPD-MA are significantly more complex and time-consuming than an analysis on aggregate data, and require significant emphasis on planning and methodology. Objective and aims. Little has been published on methodological conduct of IPD-MA completed to date. A detailed understanding of past and current practice will aid in identifying gaps and limitations. We aim to provide the most complete overview to date of IPD-MA practice, both historical and current. This will assist in creating a set of key recommendations for researchers wishing to employ IPD-MA in their work. Methods. We systematically searched MEDLINE, Embase and CDSR to identify all published IPD-MAs of randomised controlled trials up to February 2024. From these publications, we are extracting general characteristics using a purpose-built form, e.g. publication year, subject area, journal impact, geographical location, data obtained, design. Analysis will be descriptive and performed in R. Results will help us to understand the landscape of IPD-MA and changes over time.