AIMOS conference abstracts 2023

Information correct as at: 07 November, 2023. Abstracts shown in date/time order.

Tuesday 21 November

Lightning talks 1 - Meta-research

Name(s): Zuduo Zheng (In person)

Title: Advancing Reproducible Research in Transportation Studies: Insights from Personal Experience and Key Takeaways

Abstract: Reproducible research serves as a pivotal pillar of the open science movement. Within numerous Engineering disciplines including Transportation, the pervasive influence of computational analysis touches every facet of research. Yet, as algorithms grow increasingly intricate, data-driven, and sophisticated, many computational analyses have evolved into complex 'black boxes.' This opaqueness, whether intentional or unintentional, presents formidable challenges to ensuring research reproducibility. I've been promoting reproducible research in Transportation since 2015, sharing it through a journal paper, invited talks, and my website. I led a "Call for Papers" and organized a dedicated session on this topic at the 2023 Transportation Research Board (TRB) conference - the premier conference in our field, attracting over 12,000 attendees. Now, I'm planning a workshop for TRB2024 involving over 8 committees and launching another 'Call for Papers' for a journal. In my talk, I will share insights from my personal experience of promoting reproducible research and lessons applicable beyond my discipline.

Name(s): Jason Chin (In person)

Title: Metaresearch meets the justice system in the pardon of Kathleen Folbigg

Abstract: After 20 years in prison, Kathleen Folbigg was pardoned for the killing of her four children. The media narrative is that her pardon was driven by a better understanding of a genetic mutation shared by some of Folbigg's children. Several themes in metaresearch cast the Folbigg case in a different light. In fact, Folbigg's conviction was always unsafe because medical expert witnesses used undisclosed flexibility to hypothesise after the results were known. I will suggest that preregistration would have helped demonstrate to the court the fragility of these opinions. Kathleen Folbigg's pardon also raises questions about how the legal system should handle errors. Just as with scientific findings that are irreplicable, I will argue that the legal system should develop self-corrective mechanisms in order to earn its credibility.

Name(s): Kylie Hunter (In person)

Title: Responsibilities for receiving and using individual participant data from trials in health research

Abstract: While there is strong in-principle support for data sharing, in practice data are often difficult to find, access and re-use. More guidance is needed on the responsibilities of data recipients for re-using individual participant data (IPD) from trials. At AIMOS 2021, we conducted a two-hour focus group to assess views on the responsibilities of recipients of IPD, and to explore how these responsibilities could be met. This was attended by 16 conference delegates from diverse backgrounds, who discussed three common data sharing scenarios: evidence synthesis, study reproducibility and secondary analyses. Analyses revealed four themes that were synthesised into recommendations for IPD recipients toward the primary study participants (whose data are being shared) and the research team that conducted the primary study (data provider). The 'privacy and ethics' theme described the need for data recipients to prioritise the protection of participant privacy, and the recommendation to proactively share a secure data management plan and evidence of ethics oversight

with the data provider. The 'capability and resourcing' theme required recipients to demonstrate sufficient capacity to process and analyse data. The 'recognition and collaboration' theme asserted the responsibility to acknowledge the contributions of data providers and invite them to contribute to the secondary project. Lastly, the 'compliance' theme focused on the responsibility to adhere to local data sharing regulations. Implementation of these responsibilities in practice could facilitate increased data sharing.

Name(s): Olmo van den Akker (In person)

Title: The transparency and ethics of studies reusing health data

Abstract: Health data has become increasingly accessible to researchers with the advent of large databases providing routine patient data from electronic health records (e.g., Clinical Practice Research Datalink in the UK and My Health Record in Australia). The increasing accessibility of health data is promising as it allows researchers to explore or test research questions without incurring the costs and efforts of patient recruitment and data collection that are needed for randomized controlled trials. However, the guidance that researchers have to improve the transparency (i.e. registration, reporting, and data/code sharing) of studies using existing health data seems to be limited. In this study, we assess the existing literature aimed at improving the transparency of studies using routinely collected health data, and assess to what extent large patient databases ensure research transparency in practice.

Name(s): Taya Collyer (In person)

Title: What 'Strong Program' Sociology of Knowledge can do for you (. . . and Meta-Research)

Abstract: Since its proposition in the second half of the 20th Century, 'Strong Program' sociology of knowledge has been controversial, and is now rather out of fashion as a basis for empirical work in science studies. In this brief talk I advocate for a fresh look at this scholarly tradition. I introduce the Strong Program perspective and outline the diverse ways in which this view of scientific activity can challenge us to understand our own assumptions and beliefs about science, ask different kinds of questions, conduct different kinds of studies, and consider the cultural aspects of science more deeply. For those who find the approach unattractive, substantial value lies in understanding historical tensions between 'strong' and 'weak' programmes in the sociology of knowledge, as these offer relevant lessons regarding the present day challenges facing meta-research and open science.

Lightning talks 2 - Journals, writing and careers

Name(s): Natalia Gonzalez Bohorquez (In person)

Title: Altruistic versus personally rewarding attributes: the trade-offs researchers are willing to make when publishing journal articles

Abstract: Researchers have to make important decisions regarding where to publish based on several factors including journal rankings, timeliness, article processing charges, quality of reviews, target audience, topic fit, formatting, and author position. Trade-offs are made between altruistic and personally rewarding attributes, for instance, a low-ranked journal with open access, or high-ranked journal that is paywalled. Our mixed-method research explored the trade-offs researchers make when choosing where to submit papers. We used two focus groups of 18 researchers using an adapted nominal group technique, and two in-depth interviews. The analysis used an adapted version of Framework analysis. Altruistic values were found to have the least importance in the ranking exercises. The attributes identified as most relevant were ease of submission, impact factor, journal prestige and reputation, journal decision speed, quality of reviews, impact of evidence, open access, and completeness of the work. Researchers were willing to make unethical trade-offs, such as cutting analyses that were relevant to the overall story of the papers. Other factors included the role of rejection and previous experiences with journals, a strong unwillingness to publish in predatory journals, the irrelevance of cost, and the shared belief that the measure of the journal reflects the excellence of the researcher.

Name(s): Reese Richardson (Virtual)

Title: Journal hopping by research paper mills after a preferred journal is de-indexed

Abstract: Maintaining the integrity of the scientific record is of paramount importance to the public trust in science, the usage of scientific insights in law and policy, and the efficient pursuit of further scholarship. The de-indexing of a scientific journal from indices like Web of Science, Scopus and MEDLINE is a rare emergency measure if more moderate measures to ensure integrity fail. Here, we describe a new phenomenon we term "journal hopping" which undermines this measure. We directly observed and corresponded with a paper mill operator with a large and rapidly changing 'menu' of journals in which clients' manuscripts could be published. Several journals on this menu were removed and replaced by other journals once they were de-indexed. An analysis of growth dynamics at journals known to be compromised by paper mills suggests that journal hopping is a widespread industry practice. Our findings hold several practical considerations on the maintenance of the integrity of the scientific record and the deterrence of paper mill activity.

Name(s): Calvin Isch (Virtual)

Title: Narrative License in the Social Sciences: An Exploration with GPT-4.0

Abstract: Accurate scientific communication plays a pivotal role in the advancement of knowledge and the dissemination of discoveries to the wider community. However, in pursuit of recognition and impact, researchers may employ Narrative License, that is they may assert claims that are not sufficiently supported by the empirical evidence they present. While engaging and persuasive, narrative license can inadvertently lead to the overselling of scientific findings. Consequently, readers may come to false conclusions about the contributions of a study. This paper explores the phenomenon of narrative license in academic writing and the consequences it may have on the integrity of scientific research. Through the use LLMs, we assess the extent of narrative license in academic papers. We utilize broad, whole-paper-level measures of narrative license and also search for specific "rhetorical sins" including things such as overgeneralization of applicable contexts, overclaiming of impact, and understating limitations. We hope to show early results from these experiments and garner feedback on how to best measure the prevalence and impact of these concepts.

Name(s): Julia Bottesini (Virtual)

Title: Presenting JEDI: The Journal Editors Discussion Interface

Abstract: The Journal Editors Discussion Interface (JEDI) is an online community of journal editors, data professionals, and metascientists. JEDI provides an online forum that has hosted numerous discussion threads on open science and journal editing. Drawing on information and materials provided in these discussions by its members, JEDI has begun to compile a substantial collection of open science resources.

JEDI is an initiative of the Data Preservation Alliance for the Social Sciences (Data-PASS), a voluntary partnership of organizations created to archive, catalog, and preserve social science research data. JEDI's main focus is on the aspects of the editorial process that concern data and code and their management, citation, and accessibility; additional aspects of research transparency; and reproducibility, replication, and verification.

Today, JEDI is a thriving community of 430+ current, former, and incoming social science journal editors, as well as scholarly knowledge builders — data professionals and open science and metascience experts with an interest in editing and scientific publishing. In this short talk, I will present JEDI and its main components, and give a brief overview of how similar communities can help foster transparency and openness in science.

Name(s): Tony Lian (In person)

Title: Researcher adherence to data sharing policies of journals

Abstract: Background: Researchers demonstrate high in-principle support for data sharing but low level of willingness in-practice to share data from their own studies. Journals influence data sharing through their policies. We aimed to investigate researcher adherence to journal data sharing policies.

Methods: We conducted a cross-sectional analysis of original research published in 2022 in high-impact clinical medicine journals with a data sharing policy either recommending or requiring data sharing. We analysed study characteristics, initial data sharing protocols in clinical trial registries, and final data sharing plans in published data sharing statements.

Results: We included 3090 original studies in 25 journals. Most studies were publicly funded, only few commercially (n=112, 4%), Research protocols were publicly available for 10% (n=307), of which 171/307 (56%) included a data sharing statement; of these 41/171 (24%) intended to share data. Within the 68% journals recommending data sharing, 993/2028 (49%) studies included published data sharing statements, of which 868/993 (87%) intended to share data. Within the 8/25 (32%) journals requiring data sharing, only 752/1062 (71%) original studies complied and stated intention to share data.

Conclusion: Preliminary results suggest researcher adherence to data sharing requirements of journals is suboptimal and data sharing is not sufficiently enforced by journals.

Mini-notes - Sports Science

Chair: David Borg

Title: Sports science

Abstract: Kristin Sainani (Stanford University) Upping our game: How can we improve the quality of research in sports science? Methodologic errors are alarmingly common in sports science. In this talk, I will review specific examples of conspicuous statistical errors in published sports science studies and take a closer look at what these errors tell us about quality control in the field. I will also discus practical solutions. Andrew Vigotsky (Northwestern University) Error Detection in Sports Science: A Case Study. Science is said to be self-correcting, but this is only true if people put in the time and effort to correct it. What does the process of correcting science" " actually look like? In this talk, I will discuss how my colleagues and I (i) uncovered improbable data patterns in published papers in the sports science literature and the subsequent steps we took to (ii) understand its provenance and (iii) ensure the research record was correct. Specifically, (i) involves formalizing and quantifying data improbabilities, which is necessary not only to assure readers that data are, in fact, improbable, but to understand the degree of improbability. Once these improbabilities are quantified, (ii) involves carefully going through the proper channels to understand how the improbable data may have arisen. If the data's provenance cannot be ascertained, suggesting the data may not be trustworthy, (iii) necessitates reporting through the proper channels. This can be a months- or years-long process, which may not even result in correcting the scientific record. This case study serves as a real-world example of (mostly) successful error correction, which is, unfortunately, an arduous and disincentivized task. Paolo Menaspa (Australian Institute of Sport) Supporting quality research, the AIS approach. The Australian Institute of Sport is trying to address research quality issues in sport science through a dual approach that encompasses both retrospective and prospective lenses. As an example, we are considering the potential of contemporary technologies to streamline and enhance the evaluation of previously published work. Also, drawing from collaborative science approaches, we are testing concepts from 'big teams' science and examining our interaction with centralised data repositories or federated data models. Furthermore, we are investigating ways to improve data accessibility, sharing, and connection through mediums such as synthetic data generation. Karim Khan (University of British Columbia)'Consensus' in sports science/sports medicine: Shining light or crumbling edifice? Clinical sports medicine practice and sport science advice aims claims to be increasingly anchored in evidence. The term evidence-based medicine is often given a 1996 cite - Professor David Sackett's paper outlining 'what it is and what it isn't' (over 21,000 cites not that we at AIMOS pay too much attention to cites!). Evidence synthesis includes systematic reviews, meta-analysis and then the integration of these into a range of products that include 'consensus statements' and 'clinical practice guidelines'. These products sound like a good idea - who doesn't love the idea of consensus to provide guidance? However, as is always the case in research, our products are only as good as our methods. In this mini-note I will highlight the methods that have been used to develop 20-years of guiding documents for clinicians who manage patients with knee pain (manuscript has a revise and resubmit). I'll use the individual condition (knee pain) to raise 4 questions that apply to consensus statements broadly. 1. Who is in the room to decide on consensus - and are they the best people? 2. What is the agreed body of work from which to build consensus? 3. What broad process is being used to obtain consensus? (e.g. Delphi, RAND UCLA etc) 4. How is agreement defined? I will share a key pre-print of work in this domain (reporting consensus methods) and point to the ACCORD group's work. My goal is to raise awareness of an issue Dr Paul Blazey (and others) have published on and which I suspect is relevant to many colleagues at AIMOS.

Mini-notes - Scientific Whistleblowers

Chair: Simine Vazire

Title: Scientific Whistleblowers

Abstract: Eugenie Reich (Eugenie Reich Law LLC) Whistleblowing bad science and scientific fraud. Research institutions, journals, and individual scientists have financial, reputational and legal interests in obstructing self-correction and reproduction efforts. This talk will describe common experiences among whistleblowers and critics who take on such interests, and propose steps they and the community can take to facilitate self-correction and reproduction without unfairly or excessively harming individuals whose career progress is bound up with unsupported scientific claims. Kate Christian (University of Sydney) Bullying and questionable research practices: the high tolls of an overly competitive environment in academic science. In the academic science environment in Australia we have identified unacceptable levels of undesirable workplace characteristics including bullying and harassment and questionable research practices. The perpetrators of these practices are most often the supervisors or another senior colleague, people in a position of power over the victims. I will share stories of the impact on early career researchers and on whistleblowers, both from my own experience and from my research, and call for the establishment of an Australian Office of Academic Research Integrity. A/Prof Raffaella Demichelis (on behalf of the EMCR Forum of the Australian Academy of Science) The EMCR Forum: 12 years of bringing diverse voices to the decision-making tables. This presentation will introduce the EMCR Forum - who they are and what they do - and will showcase some of the Forum's advocacy goals and achievements on the need to redefine success, career structures and workload models. Bringing the voice of EMCRs and the diversity of their voices to the tables where decisions are made is a delicate and precarious task. However, a national network that has no direct link with its members' employers is by nature a much safer and freer space than a regular workplace, and the support of an organisation like the Australian Academy of Science can empower the network to achieve major milestones and amplify their voice. With this presentation, the EMCR Forum hope to initiate discussion with the audience, welcoming ideas and feedback on future advocacy goals and approaches. Liam Mannix (The Age and The Sydney Morning Herald) Scientific Whistleblowing: A How-To Guide. Liam is the national science reporter for The Age and The Sydney Morning Herald, where he regularly speaks to researchers who are trying to 'blow the whistle' on dodgy data & bad practices. He will run through some personal cases, look at the incentives to do the wrong thing and the barriers to reporting it, and then offer a few tips on how researchers can work with journalists to expose problems in science - safely.

Plenary - Clinical risk prediction models

Name: Laure Wynants (Chair: Nicole White)

Title: Clinical risk prediction models: benefiting clinical practice or piling up research waste?

Abstract: Clinical risk prediction models predict the presence of a specific disease or the occurrence of a future health-related event. They are data-driven tools based on statistical or machine learning models, designed to help healthcare professionals or patients with clinical decision-making. The typical methodology to develop and assess prediction models consists of consecutive research phases: development, internal and external validation of predictive performance, assessment of clinical impact (e.g. in RCTs), and finally implementation in care. With the rising popularity of AI, many expect clinical risk prediction models will also thrive. In this presentation, we scrutinize the current state of clinical prediction research. We will draw on the findings of two large-scale literature reviews. In the first review, we conducted a broad systematic search of bibliographic databases to estimate the number of and rise in prediction model development publications over time. A random subsample of these is followed up longitudinally to quantify how often model validation, impact assessment, and implementation are carried out. In the second review, we zoom in on the prognosis of covid-19. We will see how failure to consider the medical context in which the model will be applied, and cutting corners methodologically, make most developed prediction models unfit for clinical practice. Finally, we offer recommendations to advance the field.

Parallel sessions

Category: Discussion group

Name(s): Tom Honeyman, Nichola Burton

Title: Co-designing Future Digital Research Infrastructure

Abstract: The Australian Research Data Commons is working to identify the needs of researchers to build national digital research infrastructure solutions using a process of co-design, centring on the lived experience of Researchers. As a key voice in co-design, we are interested in the unique insights of meta-researchers in understanding both their own challenges and practices as well as those of their peers.

The ARDC is a national digital research infrastructure facility funded by the Australian Government. We partner with the research community and industry to build leading-edge digital infrastructure to support the needs of Australian researchers.

We are embarking on a new model for developing infrastructure, working through co-design to meet the needs of researchers within particular thematic areas of research–what we call the Thematic Research Data Commons. Three themes are being developed: Health and Medical Research Earth and Environmental Science Humanities, Arts, Social Sciences and Indigenous Research

In this discussion session, we would like to begin to draw out what perspective meta-researchers and open science advocates could bring to the development of Thematic Research Data Commons. We will surface the needs and challenges that you see in research across the three themes, and invite you on a journey to contribute to future co-design with the ARDC.

Category: Hackathon

Name(s): Robert Turnbull and Mar Quiroga

Title: Large Language Model (LLM) Hackathon using Structured PubMed XML Data

Abstract: We take a dataset of over 60,000 article entries in the PubMed database which have been identified as engaging with Bayesian Inference in some way. In this hackathon, we will introduce the use of LangChain to use different Large Language Models (LLMs) to ask questions of these article entries. We will experiment by writing a variety of prompts together to explore the dataset and output the results in different formats. The hackathon does not require any previous experience using LLMs but basic understanding of Python programming will be beneficial.

Category: Discussion group

Name(s): Jason Chin, Alex Holcombe

Title: Meta-Research Open Review: MetaROR, the new platform in development by AIMOS and RORI

Abstract: Meta-Research Open Review (MetaROR) is a platform in development by AIMOS and RoRI, the Research on Research Institute. One aim is to fill a gap. Meta-research is a rapidly growing field, but few to no journals are directly associated with it. MetaRoR aims to provide a free platform where meta-researchers can communicate with each other, in a place owned by scholars, one that is progressive on open research and peer review innovation. The core activity of MetaROR will be coordinating and publishing peer review of preprints related to meta-research. Traditional funding agencies also intend to work with MetaROR to improve their funding and research evaluation processes. MetaROR aims to develop a community of both authors and of journals that will be associated with the platform (similar to the Peer Community In model). After MetaROR has curated peer review of a preprint, the authors are free to submit that preprint, together with the reviews, to traditional journals. In this discussion, we will lay out MetaROR's vision in more detail and ask for input on issues such as community development, editorial policy, and how to bridge the several disciplines involved, including meta-science, philosophy of science, science studies, and policymaking.

Wednesday 22 November

Plenary - Undead theories and unstable science

Name: Stuart Ritchie (Chair: ?)

Title: Undead theories and unstable science

Abstract: Scientific research is built on the idea that evidence can clinch a question. Evidence can, in theory, support or falsify any of our ideas. But when reading the scientific literature, one finds numerous examples of questions, often important ones, that only appear to have been addressed with evidence: upon closer inspection, the studies are so low in quality, so misconceived in their design, or so outweighed by unpublished research that they provide no reliable answers whatsoever. Some authors have argued that this leaves us with a ""vast graveyard of undead theories"" - scientific ideas that remain popular, but which are impervious to falsification because of our aversion to publishing high-quality, definitive studies. Conversely, though, others have noted that the poor quality of much of the scientific literature makes it unstable and vulnerable to the sudden, drastic falsification of ideas. This includes the idea of a ""medical reversal"", where the consensus on a clinical treatment rapidly flips upon the publication of a study that's much better than anything published previously. Are these ideas in conflict? In this talk I'll discuss the many problems that arise from the proliferation of low-quality studies, and suggest ways that meta-science and open science could help us exorcise our undead theories and build a literature we can rely on.

Lightning talks 3 - Synthesis

Name(s): Jan Feld (In person)

Title: How to find phenomena in the midst of the replication crisis

Abstract: One of the main goals of the scientific enterprise is to discover phenomena, that is, stable and generalizable features of the world. The replication crisis has shaken our confidence to do so. A failed replication suggests that the original study is either wrong or the original findings are context specific. Both are concerning. In this talk, I suggest an approach to finding phenomena which has three features. First, it probes generalizability with data from many contexts and many empirical approaches. Second, it avoids publication bias by avoiding estimates that were filtered through the scientific review process. Third, it statistically accounts for sampling error using meta-analysis methods. I then apply this approach to study if men are generally more risk-seeking than women. My data covers 114 countries and 13 different measures of risk aversion. My results suggest that "men-are more risk-seeking that women in the gain domain" and "men describe themselves as more risk-seeking" are real phenomena. However, I find no systematic sex differences in risk-aversion for lotteries involving losses.

Name(s): Bill Wang (In person)

Title: Quality of randomised controlled trials, systematic reviews, and meta-analyses in paediatric surgery

Abstract: Background: There are few randomised controlled trials (RCTs) in paediatric surgery, and the risk of bias of these studies is unknown. There is also little known about the methodological quality of systematic reviews and meta-analyses in paediatric surgery. Objectives: To determine the risk of bias and reporting quality of recent RCTs, systematic reviews and meta-analyses in paediatric surgery, and the associations between these outcomes and study characteristics. Methods: This was a cross-sectional study of all RCTs, systematic reviews and meta-analyses in paediatrics surgery published in 2021. Risk of bias (as per the ROB-2 or ROBIS tools) and reporting quality (as per the CONSORT 2010 or PRISMA 2020 statements) of all studies was assessed in duplicate by two independent investigators. Risk of bias was assessed as high, unclear, or low. Reporting quality was assessed as adequate if 75% of items in the reporting guidelines were reported. Results: We identified 82 RCTs and 268 systematic reviews or meta-analyses in paediatric surgery published in 2021. More than half of RCTs (n=46, 56%) and almost all systematic reviews and meta-analyses (n=258, 96%) were at high risk of bias. Only one RCT (1%) and 4 systematic reviews and meta-analyses (1%) were adequately reported. Less than half of RCTs (n=40, 49%) and just over a quarter of systematic reviews

and meta-analyses had a registered protocol (n=77, 27%). Conclusions: Recently published RCTs, systematic reviews and meta-analyses in paediatric surgery are at a high risk of bias and have poor reporting quality.

Name(s): Annapoorani Muthiah (In person)

Title: Quality of systematic reviews and meta-analyses in dermatology

Abstract: Background: Although the number of published systematic reviews and meta-analyses in dermatology has increased over the past decade, their quality is unknown. We aimed to determine the change in the methodological and reporting quality of systematic reviews and meta-analyses in dermatology over a decade. Methods: We conducted a cross-sectional analysis of systematic reviews and meta-analyses in dermatology published in the ten highest ranked dermatology journals in 2010 and 2019. Risk of bias and methodological quality was assessed in duplicate with the ROBIS tool, and, for studies of interventions, the AMSTAR-2 tool. Reporting quality was assessed with the PRISMA 2009 and PRISMA-A 2013 statements. Results: We included 27 systematic reviews and meta-analyses published in 2010 and 127 published in 2019. There was no evidence of a difference in the proportion of studies at high/unclear risk of bias with ROBIS (Fisher's exact test = 1.00) or with critically low methodological quality using AMSTAR-2 (Fisher's exact test = 0.456), between 2010 and 2019. There was evidence of a difference in proportion of adequately reported PRISMA (t(146)=3.15, p=0.002) and PRISMA-A (t(146)=2.46, p=0.015) checklist items adequately reported between 2010 and 2019. Conclusions: No improvement was observed in the methodological quality of systematic reviews and meta-analyses in dermatology over a decade.

Name(s): Linnea Gandhi (Virtual)

Title: Research Cartography: Building a Map to Navigate and Generalize Behavioral Science

Abstract: Despite the popularity of behavioral science interventions (aka nudges), the question of their efficacy remains unresolved. Beyond issues highlighted in the Replication Crisis, such as p-hacking, the highly varied design decisions across studies make meaningful comparisons and precise inferences challenging. In this talk, we argue that the issue is not heterogeneity but incommensurability: Studies could productively differ in their designs if those differences were consistently measured. We propose a method to do so - Research Cartography - and use it to build a living map of evidence across academic and practitioner RCTs. We code each RCT (n=108 to date) across 400 dimensions - demographics, context, theories - developing an empirical language to describe, analyze, and predict study differences via machine learning. Academics and practitioners alike can use this enriched, living database to more accurately and efficiently identify where current claims are more or less precise, where they do and do not predict out of sample results, and where the field requires greater investment in evidence. We will demonstrate features of the map in the talk, such as using past RCTs to predict future results. Research Cartography as a method carries value well beyond the behavioral sciences, to a broader set of social and biomedical fields - wherever past evidence is stuck in unstructured prose. We use nudging as a timely, initial example: Recent debates questioning the generalizability of these techniques make a structured evidence map critical for the field to productively evolve.

Name(s): Yefeng Yang (In person)

Title: Robust point and variance estimation for meta-analyses with selective reporting and dependent effect sizes

Abstract: Meta-analysis produces a quantitative synthesis of evidence-based knowledge, shaping not only research trends but also policy and practices. However, meta-analytic modelling grapples with addressing two statistical issues concurrently: statistical dependence and selective reporting (e.g., publication bias). Here, we propose a two-step procedure to tackle these challenges. First, we employ bias-robust weighting schemes under the generalized least square estimator to obtain less biased population mean effect size estimates by mitigating selective reporting. Second, we use cluster-robust variance estimation to account for statistical dependence and reduce bias in estimating standard errors, ensuring valid statistical inference. By re-analysing 448 published meta-analyses, we show that our approach is effective at mitigating bias when estimating mean effect sizes and standard errors. To assist adoption of our approach, we provide a website showing a step-by-step tutorial. Complementing the current practice with the proposed method can facilitate a transition to a more pluralistic approach in quantitative evidence synthesis.

Parallel sessions

Category: Discussion group

Name(s): Maria del Mar Quiroga (plus Prof David Goodman)

Title: Strengthening digital research literacies in the Humanities, Arts, and Social Sciences

Abstract: In her 2012 book "How we think", Katharine Hayles speculated that by about now, 50% of humanities scholars would be "seriously engaged" with digital technologies (meaning something more than email, web browsing, and Microsoft Office). That clearly didn't happen. In 2022, at the University of Melbourne, we got access to funds to establish the HASS Taskforce: an initiative to help researchers in the Humanities, Arts, and Social Science to strengthen their digital research literacy through targeted, hands-on digital training and support. We ran surveys to better understand what support researchers wanted and used that to develop a program of free workshops throughout 2023. Topics have included Python, Data Visualisation, APIs, Omeka, Constellate, TLCMap, and Web scraping. Feedback on these sessions has been very positive. Based on the Carpentries philosophy, we keep a high instructor/helper to attendee ratio, which we think is key to attendees' positive experiences. However, it has not always been easy to find HASS scholars willing to attend our workshops; we have at times resorted to last-minute advertising in general graduate student forums just to get enough attendees, which comes at the cost of a loss in our target audience and high attrition rates. What are we doing right and wrong in working towards a more digitally engaged, capable, and empowered humanities, arts, and social science academy? How can we better communicate the opportunities that come with stronger digital research literacy? Whose responsibility is it to do that anyway?

Category: Workshop

Name(s): Eugenie Reich

Title: Technical Critique Workshop: venues, approaches, risks and consequences

Abstract: This workshop will tackle the options that exist in the scientific community for publicizing candid technical critiques and takedowns of flawed and fraudulent research. We will discuss communication problems and legal and ethical issues at stake in a critic or whistleblower's choice of forum: from personal email to group meetings to conferences to preprint servers to journals to the news media to social media to the courts. We will also discuss different approaches that critics can take in presenting their data or theoretical analysis and the different kinds of language that they may use to convey what they think went through. Finally we will study some real-life examples of different technical critiques (which I will share on handouts for those there in person) and talk about the content, the style of presentation, what the individual and community ramifications might have been of sending the critique into the world, and what types of follow up should happen from different parties involved. An important theme is to what extent one should just say what one thinks versus consider the trade-off with the risks and consequences of harsh critiques for individuals (both critics and targets of criticism), institutions, funders and research fields.

Lightning talks 4 - Statistics and study design

Name(s): Robert Turnbull (In person)

Title: Bayesian Inference in Medical Literature: Trends in PubMed Data

Abstract: Ronald Fisher's rejection of Bayesian Inference profoundly influenced the field of statistics for decades. However, in the past half-century, significant developments in computational techniques, the widespread availability of Bayesian software and a growing acknowledgment of the role of subjectivity in research have ushered Bayesian analysis from the margins into the scientific mainstream. This lightning talk presents the results of a comprehensive meta-study conducted on the PubMed database, comprising over 35 million entries. Our research examines the increasing trend of articles engaging with Bayesian Inference, particularly since the early 2000s. To achieve this, we used Large Language Models (LLMs) to systematically classify these articles. Our study not only traces the rise of Bayesian Inference in medical literature but also seeks to answer questions about its impact. We aim to uncover the specific areas within this corpus where Bayesian Inference has made the greatest contributions and explore the implications of these trends.

Name(s): REN ZHILIN (In person)

Title: Interventions during study design and conduct to reduce risk of bias in interventional studies

Abstract: Background: Poorly designed studies may lead to biased results that skew evidence. We reviewed publications to identify interventions during studies design and conduct to reduce risk of bias. Methods: Since 19th September 2022, we searched MEDLINE, Embase, Cochrane Library and nine grey literature sources. Keywords for the components of interventional studies were combined with phrases related to bias to form the search strategy. Publications were included if they described the implementation and effectiveness of interventions which aimed to reduce risk of bias in interventional studies. Observational studies, or publications describing an intervention undertaken within observational studies were excluded. Results: We included 2 studies after screening 36,598 publications. The first intervention, education and training for researchers during study design, included implementing a more rigorous participant screening process and systematic participant tracking program during study conduct, reducing loss to follow-up rate and missing data compared to similar studies. The second intervention, independent clinical events committee during study conduct, reduced bias due to conflicts of interest affecting the analysis and interpretation of results. Conclusion: Despite the major impact of risk of bias on evidence-based medicine, there are too few existing interventions to address this, pointing toward a major gap in the evidence base.

Name(s): David Borg (In person)

Title: Meta-Analysis Prediction Intervals are Under Reported in Sport and Exercise Medicine

Abstract: Aim: To estimate the proportion of meta-analysis studies that report a prediction interval in sports medicine; and the proportion of studies with a discrepancy between the reported confidence interval and a calculated prediction interval. Methods: We screened, at random, 1500 meta-analysis studies published between 2012 and 2022 in highly ranked sports medicine and medical journals. Articles that used a random effect meta-analysis model were included in the study. We randomly selected one meta-analysis from each article to extract data from, which included the number of estimates, the pooled effect, and the confidence and prediction interval. Results: Of the 1500 articles screened, 866 (514 from sports medicine) used a random effect model. The probability of a prediction interval being reported in sports medicine was 1.7% (95% CI = 0.9%, 3.3%). In medicine the probability was 3.9% (95% CI = 2.4%, 6.6%). A prediction interval was able to be calculated for 220 sports medicine studies. For 60% of these studies, there was a discrepancy in study findings between the reported confidence interval and the calculated prediction interval. Prediction intervals were 3.4 times wider than confidence intervals. Conclusion: Very few meta-analyses report prediction intervals and hence are prone to missing the impact of between-study heterogeneity on the overall conclusions. The widespread misinterpretation of random effect meta-analyses could mean that potentially harmful treatments, or those lacking a sufficient evidence base, are being used in practice.

Name(s): Shiva Raj Mishra (In person)

Title: Routine exclusions in nephrology trials in Australia from 2010 to 2023: A meta epidemiological study

Abstract: The study aims to determine the prevalence of routine exclusions in nephrology clinical trials in Australia registered in ANZCTR or ClinicalTrialsGov from 2000 to 2022. We counted the number of trials excluding patients of the prespecified exclusion groups (demographic, geographic, reproductive, comorbidities, disease severity, health services, language) and calculated the prevalence of exclusion for each group. Of 162 eligible clinical trials protocols. The median sample size at recruitment was 80 (IQR: 40 to 166). Nearly 74% of the trials applied at least one exclusion criteria. 34.1% and 14.6% of trials excluded pregnant and breastfeeding women. More than two thirds of trials excluded patients with at least one comorbidity: 35.8 (n=44) of the trials excluded patients with cardiovascular complications (e.g., arterial fibrillation, heart failure), severe kidney damage (including failed transplant) 34.1 (n=42), cancer 22% (n=27), liver disease 21.1% (n=26), mental health conditions 21.1(n=26), and diabetes complication 19.5% (n=24). Further, 2.4% (n=3) and 13% (n=16) of the trials excluded participants who were either “too trail” or reported having “low life expectancy” by trialists. Future studies are needed to explore discretionary exclusionary practices in clinical trials, as well as exploring framework/guidelines for promoting diversity in participant’s recruitment strategies.

Name(s): Taya Collyer (In person)

Title: The Eye of the Beholder: How do public health researchers interpret regression coefficients?

Abstract: Calls to improve statistical literacy and transparency are widespread, but empirical accounts describing how researchers understand statistical methods are lacking. This study explored variation in researchers' interpretation and understanding of regression coefficients, and extent to which these are viewed as straightforward statements about health. Method: Thematic analysis of qualitative data from 45 interviews (academics from 8 countries, representing 12 disciplines). Three concepts from sociology of knowledge and science studies aided analysis: Duhem's Paradox, the Agonistic Field, and Mechanical Objectivity. Some interviewees view regression as discovering real relationships, while others indicated models are not direct representations. Coefficients were generally not viewed as mechanically objective, instead interpretation was described as iterative, nuanced, and sometimes depending on prior understanding. Researchers consider many factors when evaluating results, including knowledge from outside the model and whether results are unexpected. Interviewees repeatedly highlighted the role of the analyst, reinforcing that it is researchers who answer questions and assign meaning, not models. Regression coefficients were not viewed as complete or authoritative statements about health. This contrasts with teaching materials presenting statistical results as straightforward representations, subject to rule-based interpretations. Attempts to influence conduct and presentation of regression models should be attuned to the myriad factors which inform their interpretation.

Parallel sessions

Category: Discussion group

Name(s): Olmo van den Akker

Title: Open peer review: Solving or exacerbating science's problems?

Abstract: Open peer review, a system in which authors and reviewers are aware of each other's identity, and review reports are published alongside the relevant article, could solve many problems in science. For example, it could yield more constructive and higher quality reviews, could alleviate the pressure on the peer review system, and may even cause the demise of predatory journals. On the other hand, concerns have been levied about biases and retaliation once author names become known. To date, not much empirical research has been conducted to assess whether the proposed benefites and concerns are valid. In this session, we will discuss the pros and cons of open peer review, and devise ways to empirically test how these pros and cons pan out in scientific practice.

Category: Discussion group

Name(s): Jonathan Williams, Anna Lene Seidler, Aidan Tan, Kristan Kang

Title: Secondary data use scenarios

Abstract: Data sharing is a core element of the Open Science movement and adds tremendous value to existing data. Yet, structured guidance around the different use-case of secondary data use, their advantages, but also risks and limitations is currently lacking. The Health Studies Australian National Data Asset (HeSANDA) program by the Australian Research Data Commons (ARDC) has recently launched national infrastructure platform to allow researchers to access and share data from health studies. ARDC now seeks to engage with researchers to define and communicate the scientific value and use cases of secondary research using clinical trials data to ensure value of this asset. As part of these activities, this discussion group aims to explore the development of a theoretical framework for the kinds of secondary research to be undertaken with clinical trials data shared on HeSANDA platform. Participants will be guided through a series of questions and scenarios, including advantages and risks/limitations of different use cases for secondary research. This will assist developments of user guides on the HeSANDA platform. Ultimately, this work will inform a roadmap for developing a health research data community whose research could be supported using HeSANDA infrastructure. Participants will have opportunities for recognised further input.

Category: Hackathon

Name(s): Beth Clarke

Title: The secret to getting metaresearch done. Creating resources for conducting undergraduate metaresearch projects

Abstract: There is a great deal of untapped potential in running metaresearch studies as undergraduate research projects. Not only do metaresearch projects provide students with the necessary skills to conduct rigorous research, but they also enable you — the supervisor — to get important and impactful metaresearch done. For these projects to reach their full potential, we would like to equip aspiring meta-research supervisors with the resources to make these studies happen. In this hackathon, we aim to create resources on topics such as: the pros and cons of conducting metaresearch as undergraduate projects (and discussing these with students); advice throughout the metaresearch cycle (e.g., identifying research questions, appropriate methods, publication etc.); supplying examples of past projects. We also foresee a need to coordinate these studies. Thus, a second aim of this hackathon is to establish a grassroots network for these projects, so that the metaresearch community can: identify research questions of high priority; solicit technical, coding, and other forms of help from one another; and avoid research waste by centralizing efforts rather than sprawling research. Our ultimate goal is to produce a paper or other output that will serve as a resource for researchers. If you have thoughts on any of those things, join us! Any background and experience level are welcome.

Mini-notes - AIMOS top-up scholarship winners

Chair: Rose O'Dea

Title: AIMOS top-up scholarship winners

Abstract: Savannah Lewis (The University of Alabama) Who/Why Big Team Science. Robin Guelimi (Université Paris Est Créteil) Overview of the redundancy and factors of variability of the overlapping systematic reviews with network meta-analysis evaluating the efficacy and safety of systemic treatments in immune-mediated inflammatory diseases. Elliot Gould (University of Melbourne) Investigating challenges and solutions to improve the reproducibility and transparency of model-based research in applied ecology and conservation decision-making Benjamin Meghreblian (Cardiff University) Registered Reports Community Feedback - a website for collecting peer review experiences from authors and reviewers

Thursday 23 November

Plenary - The impediments to high-risk, high-return research

Name: Carl Bergstrom (Chair: Losia Lagisz)

Title: The impediments to high-risk, high-return research

Abstract: Scientific researchers may be driven by curiosity, but they are constrained by the realities of the scientific ecosystems in which they operate and motivated by the incentives with which they are confronted. We can use mathematical models of the research enterprise to understand how scientific norms and institutions shape the questions we ask, the efficiency with which we work, and the discoveries we make about the world around us. In this talk I present a pair of mathematical models aimed at revealing why scientists are reluctant to propose and conduct high-risk research. In the first vignette we look at how peer review filters - ex ante review as for grant proposals and ex post review as for completed manuscripts - shape the types of questions that researchers pursue. In the second vignette, we develop an economic ""hidden action"" model to explore how the unobservability of risk and effort discourages risky research. Scientific norms and institutions are not god-given; we create and maintain them. If we can understand their consequences, we have the potential to nudge the norms and institutions in directions better tailored to our contemporary research questions and technologies.

Lightning talks 5 - Diversity

Name(s): Tatiana Chakravorti (Virtual)

Title: Awareness and Perceptions of Research Reproducibility in India: A Study with Indian Researchers

Abstract: Several large-scale projects tried to reproduce the results of published scholarly articles but failed in recent years which significantly points out a huge replication crisis in science. It was found that 70% of these researchers were not able to reproduce other's work as well as some of their own published work which creates concerns about losing public trust and credibility. In this research, we have explored the state of the reproducibility crisis and open science in India and the awareness of Indian researchers toward scientific research. The core challenges related to reproducibility research practices and how the researchers can be incentivized. 19 interviews were taken to understand the situation in India and to overcome the diversity we sent a survey where we got responses from 72 professors from the highly reputed institutes from India. According to these professors, a major change is needed for the publication process which links the Open Science Movement.

Name(s): Shiva Raj Mishra (In person)

Title: Conceptualisation, operationalisation and utilisation of equity, diversity and inclusion in clinical trials

Abstract: Ensuring equity, diversity, inclusion (EDI), helps generating highest quality evidence for interventions in the populations most likely to benefit. This study sought to understand the conceptualisation (definition), operationalisation (measuring/coding), and utilisation (analyses) of EDI in clinical trials. We systematically reviewed literature from Pubmed/Medline (via NLM) and Google Scholar from 1990 to 2023. Additionally, we searched a sample of websites of health actors (n=43 'actors') across the research lifecycle to identify literature not captured by database searches. We reviewed 2385 titles/abstracts and 43 'related:URL' searches, included 72 (3.0%) in analyses. Studies delineated EDI as interconnected concepts rather than distinct constructs. These concepts were intertwined and were often reinforcing. For example, efforts to enhance diversity may also promote equity and foster inclusion.Several frameworks, tools and metrics were identified for EDI assessment across the research lifecycle. All publishers (6/6) provided a statement representing EDI; followed by 2/3 trial registries, 8/13 research institutions, 7/12 journals, and none of ethics committee and data repositories reported statement on EDI. Our study also revealed that the proportion of patients excluded varied widely among different exclusion criteria. Future research could explore the impact of different EDI criteria on trial outcomes and the generalisability of trial results.

Name(s): Totoro Nakagawa-Lagisz, Kohaku Nakagawa-Lagisz (In person)

Title: Dead Scientists in Nature: Piloting Human-AI Collaboration for Evidence Synthesis

Abstract: When prominent scientists die, they get their final recognition in the form of published obituaries. From scientific obituaries we can see which researchers are honoured and how the obituaries describe them. But unfair writing about women in obituaries can make the community view women differently than men in science. Thus, it is important to assess if obituaries of men and women are written differently.

This project aims to survey the number of obituaries of 30 male and 30 female scientists in the journal "Nature" from 2011 to 2021. We analyse the way that these past scientists are portrayed in these obituaries using the "Finkbeiner Test", which notes excessive emphasis on gender and personal circumstances. As such, this work provides evidence on gender-related biases in scientific recognition.

We consider the relationship between different combinations of gender of the obituary authors and gender of the deceased scientist. Further, this pilot project is conducted using in parallel both a human researcher and a custom-designed Large Language Model bot. This allows us to compare these two approaches and test how they can be best combined for faster and more reliable information extraction from the text for future meta-research projects.

Name(s): Kanghui WEI (In person)

Title: The Association of Women Authors with Women Enrolment in Cardiovascular Trials in LIMCs

Abstract: Cardiovascular diseases (CVDs) are a leading cause of mortality in low- and middle-income countries (LMICs). Despite high rates of CVDs among women in these regions, there exists a significant gap in women's enrollment, as well as authorship in clinical trials. A systematic review was conducted to investigate clinical trial participation and authorship in cardiovascular clinical trials in LMICs (defined based on World Bank Classification). Data from 198 trials were included in the analysis. The overall proportion of women in authorship was 34%, while the proportion of women patients enrolled in these trials was 36%. Women's representation in authorship overall was positively but not statistically associated with women's participation in cardiovascular clinical trials (P-value =0.921). The top three countries with the lowest enrollment in clinical trials and lowest representation of women in authorship were Egypt, India, Bangladesh. Contrarily, Uganda, Tunisia, and Pakistan had overall the highest women enrolment and highest representation of them in authorship. Relationships were similar when using representation in first, and last (senior) authorship compared to any authorship positions. Our study found gender gaps in trial participation as well as authorship practices in LMICs. Efforts are needed to ensure gender-inclusive practices in clinical trials.

Name(s): Malgorzata Lagisz (In person)

Title: What can we learn from Best Paper awards across disciplines?

Abstract: Research awards provide recognition to outstanding researchers for excellent research (at least that is what they usually claim). But what "excellent research" means? Can we tell what the assessors and awarding bodies value from award descriptions? Do they value robust, transparent, and inclusive research? Do they incentivise adherence to core practices of Open Science? And, who gets these awards?

In this talk, I will present results of a collaborative project evaluating 222 Best Paper awards across disciplines. Notably, this work has been initiated during an AIMOS 2022 conference hackathon – so it is a perfect time and place to present our findings.

We found that journals and learned societies administering Best Paper awards publicly provide very little detail about their awards. Award descriptions almost never mentioned concepts that align with Open Science. Between 2001 and 2022, most individual winners were USA-affiliated, while researchers from the Global South and developing countries were uncommon. Sixty one percent of individual winners were men.

Overall, Best Paper awards miss the global calls for providing greater transparency and equitability in science. We can see them as an untapped opportunity for aligning incentives in the publication process and research recognition with values for greater openness and inclusivity.

Mini-notes - Open data

Chair: Anna Lene Seidler

Title: Open data

Abstract: Kyle A Sheldrick : Are meta-analyses and guidelines based on summary data alone inherently unreliable? Aidan Tan : Analysis of data sharing policies, barriers and facilitators across the research life cycle. Kristan Kang : Setting up data sharing infrastructure in Australia. Tom Hardwicke : Open science across disciplines.

Lightning talks 6 - Errors and improvements

Name(s): Kabir Manandhar Shrestha (In person)

Title: AI in Interdisciplinary Research: Bridging Domains and Advancing Open Science

Abstract: The rapid evolution of Artificial Intelligence (AI) offers transformative avenues for metaresearch and open science. As diverse disciplines intersect with AI, this discussion seeks to inspire researchers to harness AI's potential, enhancing both their domain and AI itself.

Benefits of AI in Interdisciplinary Research:

1) Enhancing Traditional Research: Incorporating AI into various domains introduces innovative methodologies, leading to novel outcomes and broadening research perspectives, thereby contributing to metaresearch.

2) Refining AI through Diverse Data: Subjecting AI to unique datasets from varied disciplines not only challenges but also refines AI methodologies, ensuring continuous improvement.

3) Elevating Open Science: The fusion of AI with non-technical fields necessitates open science approaches, giving birth to insightful websites, comprehensive data repositories, and open-source code, fortifying research transparency and collaboration.

Case Studies as Evidence:

1) Political Rhetoric Project: This collaboration employed Natural Language Processing to analyze Malcolm Fraser's speeches, revealing evolving political themes. The insights, presented via an open science platform, offer a unique lens into political rhetoric evolution.

2) Savannah Land Cover Forecasting: By harnessing deep learning, this collaboration predicts future land coverage in Northern Australia's Savannah. The findings, supported by open-sourced code, have profound implications for environmental conservation and research transparency.

Name(s): David Smailes (Virtual)

Title: A citation analysis of a flawed meta-analysis of the efficacy of CBT for bipolar disorder

Abstract: It is often argued that one of the key strengths of science is that it engages in 'self-correction'. However, there is some evidence to suggest that science does not self-correct very effectively – for example when looking at citation patterns of studies that have failed to replicate. Here, we aimed to examine another way in which science might be expected to self-correct. We identified a meta-analysis that contained a clear and obvious error, which – when adjusted for – substantially changed the conclusions that could be drawn from the meta-analysis. We then examined how that meta-analysis had been cited, coding the ways in which 112 publications discussed/used the findings from the meta-analysis. We rated whether they had discussed the meta-analysis' findings in positive terms (i.e., uncritically noting that the meta-analysis supported the efficacy of CBT for bipolar disorder), in neutral terms, or critically/negatively (e.g., noting that the conclusions made by Chiang et al. may have been incorrect). Only one of the publications we coded raised concerns about the accuracy of the conclusions made in in the meta-analysis, with the vast majority of studies discussing the findings positively/uncritically. Our findings are consistent with previous studies that have suggested that science typically does not 'self-correct' in an effective and timely manner.

Name(s): Paulina Stehlik (In person)

Title: ENHANCE survey: Medical specialty training research requirements contribute to research waste.

Abstract: Background Our work has suggested that College-set research requirements for medical specialty trainees focus on leading research rather than research skills development. We hypothesised that this leads to research waste. Methods We conducted an online survey of current and past trainees across Australia and New Zealand medical specialty colleges to investigate their mandated research experiences and outputs. Participants also submitted projects for reporting and design quality evaluation. Results 372 trainees from all 16 major colleges participated, 177 had completed a project. 41% generated a research question on their own, 26% did not conduct a literature review before starting, and 40% did not have a publicly available protocol. 38% agreed that they had the knowledge and skills to conduct research, 19% for seminar and 56% for methodological expertise access, and 67% for adequate supervisor support. 52% of the projects were designed by the trainee alone, and 49% remain unpublished. Of the 29 evaluated project uploads, most had significant gaps in reporting quality. 28 had moderate-to-high risk of bias. Qualitative responses suggest trainees felt projects took up too much time and contribute to research waste. Conclusion Current college research requirements see trainees conducting "poor quality... box-ticking research" in suboptimal conditions, contributing to research waste. Colleges should allow trainees to take on smaller roles in larger projects to focus on skills development, with well-resourced pathways for the few who wish to become leaders in research.

Name(s): Danielle Oste (In person)

Title: Non-verifiable cell lines in cancer research papers describing human gene research

Abstract: Reproducible laboratory research relies on correctly identified reagents. We have described preclinical human gene research papers published between 2008-2019 that described wrongly identified nucleotide sequence reagent(s), including papers that studied the human miR-145 gene. Manually verifying reagent identities in more recent miR-145 papers from 2020-2022 found that most miR-145 papers described the use of wrongly identified nucleotide sequence(s) and/or cross-contaminated human cell line(s). We also found 5 cell line identifiers (BGC803, BSG803, BSG823, GSE1, TIE-3) in miR-145 papers and a further 6 identifiers in other papers that do not correspond to known human cell lines. While some of these occurrences likely reflect misspellings of known cross-contaminated cell line identifiers, to date we have also found publications that refer to BGC803, BSG823 or MGC823 as independent cell lines. We have not been able to find publications that describe how these 3 cell lines were established, and these cell lines do not appear to be indexed in claimed cell line repositories with external catalogues. While some publications stated that short tandem repeat (STR) profiles were generated for BGC803 or MGC823 cell lines, no STR profiles could be identified. In summary, non-verifiable cell lines represent challenges to research reproducibility and require further investigation to clarify their identities.

Name(s): David Wilkinson (In person)

Title: Software Citation Taxonomy (SofCiT): A protocol for appropriate software citation

Abstract: As technology progresses at ever accelerating rates it allows for bigger and better things to be done in science. Technological advances are allowing for new data collection methods while computational advances are letting us analyse bigger and more complex datasets. Research is becoming increasingly software dependent over time but the machinery of academia is failing to keep pace. The burden of software development usually falls to early career researchers, graduate students, or embedded professional staff, but this work isn't given the equivalent recognition to traditional primary research outputs like publications for career progression. Software also plays an important role in the "replicability crisis", where the lack of transparency in software reporting in publications creates an additional hurdle to reproducing published results.

We're proposing a new software citation protocol for publications called Software Citation Taxonomy (SofCiT) to address these issues. The protocol builds off of the Contributor Role Taxonomy (CRediT) and Methods Reporting with Initials for Transparency (MeRIT) frameworks to to standardise the way software is cited in the literature. This includes several core tenets such as: a) a Software Statement (ala CRediT) to list what software as used and how, b) weaving critical software into the Methods section directly (ala MeRIT), and c) appropriate software citation requires versions.

Parallel sessions

Category: Discussion group

Name(s): Cooper Smout

Title: WisdOHM: A novel app and metascientific system for ranking and rewarding diverse contributions

Abstract: Numerical evaluation has the potential to revolutionise academia, but adoption of such systems remains low due to misaligned incentives.

We will open this discussion circle with a round of introductions and any relevant experience using numerical evaluation in academia. I will then introduce WisdOHM, a novel open source app and metascientific system designed to rank and reward community contributions, thus overcoming the 'incentive problem' by directly rewarding contributors.

I'll give a brief history of its development under the auspices of our charitable organisation, Open Heart + Mind (OHM), including the multiple versions we've prototyped before settling on a simple, pairwise comparison protocol that is user-friendly, robust, accessible and amenable to metascientific analysis.

Participants will then be invited to login to the app and review contributions from our recent contribution-based event, by way of analogy for reviewing research contributions. We will then open up a round of Q&A, to clarify anything that is unclear.

I'll then invite a final discussion round about potential use cases in academia, with a focus on:

(a) platforms or projects working on similar problems, so we can collaborate instead of compete
(b) types of contributions we could review (e.g., papers, code, data, conference workshops)
(c) dimensions we should be interested in (e.g., reliability)

The outcomes of this discussion group will have real-world implications in how we develop the app. I invite a lively discussion with anyone interested in the future possibilities of academic evaluation.

Category: Workshop

Name(s): Wentao Li

Title: Assessing trustworthiness of randomised controlled trials with and without individual-level data

Abstract: Background: The inclusion of untrustworthy randomized controlled trials (RCTs) in systematic reviews and meta-analyses is a growing concern that troubles producers and consumers of evidence synthesis. There is currently limited guidance on how to evaluate trustworthiness in RCTs.

Objectives: 1. To provide an overview of methods to identify untrustworthy RCTs; 2. To gain hands-on experience in how to apply simple and effective methods that identify untrustworthy RCTs; 3. To discuss options if evidence is found that suggests poor trustworthiness in RCTs.

Description: This workshop will consist of a brief introduction to the rationale for assessing trustworthiness in RCTs and an overview of available methods that assess trustworthiness. Then, there will be two interactive sessions on the practice of selected methods that identify problems in RCTs with and without the availability of individual participant data, followed by a short discussion on how to handle RCTs identified as not trustworthy.

Planned agenda: Workshop 1 I. Introduction to trustworthiness issues in RCTs (10 minutes) II. Overview of methods that could be used to identify trustworthiness issues in RCTs (20 minutes) III. Practice 1: use the TRACT checklist to identify problems in RCTs without access to individual-level data (60 minutes) Workshop 2 IV. Practice 2: use the CHIPPR toolkit to identify problems in RCTs with individual-level data (70 minutes) V. Discussion on how to handle RCTs identified as not trustworthy (20 minutes)