Presentations 1: Planning, collaboration and review management

Authors: Wagner, Gerit; Prester, Julian
Title: CoLRev: A pipeline for collaborative and git-based literature reviews
Abstract: Conducting highly collaborative literature reviews remains a key challenge and compelling visions for pipelines covering the literature review process end-to-end, i.e., from problem formulation to write-up, have yet to be proposed. We contend that, similar to other reproducible research contexts, git offers a viable data management foundation for literature reviews. However, the unique characteristics of literature reviews have yet to be fully considered, and corresponding design principles are yet to be proposed. In this context, our work builds on years of iterative prototype development, evaluation, and refinement. The proposed talk focuses on the aforementioned challenges and proposes potential solutions. Our objectives are twofold: First, we aim to demonstrate a novel data management and tool pipeline (CoLRev), and second, we summarize its key principles. Integrating such a pipeline with other emerging tools, we see exciting opportunities for evidence synthesis communities to transform the conduct of collaborative literature reviews.

Authors: Diakonidis, Theodoros
Title: screenmedR: a package for automatizing the screening of publications for meta-analysis or systematic reviews, using pubmed database.
Abstract: A new program is introduced, in R computing language, fast and accurate, that could save researcher's time in sifting the publications coming from a Pubmed search, in order to find the relevant ones for a systematic review or a meta-analysis. Suppling an input of 4-5 publications the researcher believes that belong to its study, the program and most specifically the function screenmed() can diminish its initial pubmed search 60% to 80% with almost 100% accuracy. The program uses the abstracts of the publications and apply text mining methods in combination with hierarchical clustering, an unsupervised machine learning practice, and cosine similarity to find similarities among them. It also provides 2 functions which use mesh terms definitions to find similarities between publications. One mesh_clean_bq()to find common mesh terms between two groups of publications and the second mesh_by_name_bq()to find specific mesh terms from a group of publications.

Authors: Nevill, Clareece; Quinn, Terry; Cooper, Nicola; Sutton, Alex
Title: MetaImpact: Designing future studies whilst considering the totality of current evidence
Abstract: When designing new studies, there is a balancing act regarding how many people to recruit. Too few, the trial may not detect an effect; too many, some participants would have undergone inferior treatments unnecessarily. Both scenarios are wasteful and unethical. In the UK, governing bodies must approve new treatments before offering them to the public. To aid decision-making, systematic reviews and meta-analyses are often presented where all relevant evidence is systematically found and combined to give an overall picture. Therefore, new trials should ideally add to the current evidence and influence a future review. Sutton et al developed a method for estimating the sample size of a new trial such that it influences a future review. This involves simulating a new trial using parameters from the current meta-analysis, adding it to the review, and then seeing how the results changed. Repeating this multiple times estimates the 'power' of the sample size – the proportion of

simulations that give the desired effect. This project aimed to create an interactive web-app for researchers to easily utilise these methods themselves. Using software 'R' and 'shiny', a free web-app was created, MetaImpact, to estimate the power of a future study with a certain sample size having an impact on a review. Educational features range from information boxes explaining different elements of the calculator, to plots illustrating how the estimation is calculated. Past reviews were utilised to assess the benefit of MetaImpact by removing the most recent addition, adding a new trial generated using MetaImpact to the review, and then comparing the result to the original. MetaImpact has potential to benefit patients and research by encouraging ethical sample sizes and reducing 'wasteful' trials.

Authors: Vembye, Mikkel
Title: Conduction Power Analysis in Meta-Analysis With Dependent Effect Sizes Using the POMADE R package
Abstract: Meta-analytic models for dependent effect sizes have grown increasingly sophisticated, which has created challenges for a priori power calculations. In a recent paper (https://doi.org/10.3102/10769986221127379), my colleagues and I introduced power approximations for tests of average effect sizes based upon several common approaches for handling dependent effect sizes. In this presentation, I will provide an overview of the POMADE R package, which implements these methods. The package includes functions to conduct power analysis based on the correlated-and-hierarchical effects (CHE) model, as well as Multi-Level Meta-Analysis (MLMA) and the correlated effects (CE) models. I will focus on how to approximate power for the CHE model, varying assumptions regarding the approach to variance estimation and the degrees of freedom used in power formulas, as well as the assumptions regarding the precision of primary study effect size estimates and the number of effect sizes per study. In addition to estimating power, I will demonstrate R functions for approximating the minimum detectable effect size and the number of studies needed to find a certain effect size of practical concern with pre-specified values of statistical significance and power. I will also provide rough guidelines for how plausible design factor values and model parameters can be selected and how to plot and present the given analyses.

Authors: Whaley, Paul; Wattam, Stephen; Scott, Anna Mae; Vidler, John
Title: Automated documentation of standards-compliant systematic review protocols
Abstract: Semantic labelling of data facilitates its reuse. When combined with general-purpose protocols, labelled data can enable the automatic generation of standards-compliant research documentation, without the researchers involved needing explicit knowledge of what those standards are. Here we present a proof-of-concept of how this could work, and discuss limitations in the technology stack that could be addressed by the automation community.

—

Presentations 2: Searching and record management 1
Authors: Riley, Trevor; Young, Sarah; Hair, Kaitlyn; Wallrich, Lukas
Title: CiteSource R Package & Shiny
Abstract: This presentation will review the CiteSource R package and Shiny app. CiteSourse began as part of ESMARConf2022 and gives users the ability to deduplicate reference files,

while maintaining custom metadata fields. This functionality allows users to compare search results of literature sources, search methodologies, and strategies. The package also allows users to analyze search results, methods, etc. across both title/abstract screening and full-text screening phases. This tutorial will provide an overview of the CiteSourse R package/Shiny app, and will explore various use cases through unique vignettes.

Authors: Lajeunesse, Marc
Title: A study-centric reference manager for research synthesis in R
Abstract: Reference managers like Zotoro, Mendely, and Endnote are often shoehorned for key tasks like retrieving, organizing, screening, and coding studies for systematic reviews and meta-analysis. However, the biblio-centric, spreadsheet-like UIs of these tools are less than ideal for the fastidious study-level work typically needed for research synthesis. Here I introduce an experimental R package that offers an alternative reference managing design that de-emphasizes tabular interfaces for a more study-centric UI that enhances interactivity, task tracing, coding, and reference readability. The primary goal of the software is to improve user experience and make the diverse and repetitive tasks of research synthesis more palatable.

Authors: Bougioukas, Konstantinos; Diakonidis, Theodoros; Mavromanoli, Anna; Haidich, Anna-Bettina
Title: ccaR: a package for assessing primary study overlap across systematic reviews in overviews
Abstract: An overview of reviews aims to collect, assess, and synthesize evidence from multiple systematic reviews (SRs) on a specific topic using rigorous and reproducible methods. An important methodological challenge in conducting an overview of reviews is the management of overlapping information and data due to the inclusion of the same primary studies in SRs. We present an open-source R package called ccaR (https://github.com/thdiakon/ccaR) that provides easy-to-use functions for assessing the degree of overlap of primary studies in an overview of reviews with the use of the corrected cover area (CCA) index. A worked example with and without consideration of chronological structural missingness is outlined, illustrating the simple steps involved in calculating the CCA index and creating a publication-ready heatmap. We expect ccaR to be useful for overview authors, methodologists, and reviewers who are familiar with the basics of R. We also hope our package will contribute to the discussion about different methodological approaches in implementing the CCA index. Future research could further investigate the functionality of our package and other potential uses as well as the limitations.

Authors: Twente, Leonie
Title: Implementing text mining to optimize search term selection for systematic reviews in language education: a case study
Abstract: Systematic reviews (SR) collate all available empirical findings to answer a clearly formulated question. The systematic and transparent approach minimizes the risk of biases on the selection and evaluation of relevant studies. Ideally, the search strategy finds both all documents relevant to the question ("sensitivity") and as few irrelevant documents as possible ("precision"). Identifying search terms for searches in electronic databases is a challenge, in particular for SRs in educational research where there is no standardized

system of terms such as "MeSH" in medicine. Since SRs in highly interdisciplinary fields require searching databases of different disciplines, the keywords assigned in thesauri are not an optimal solution. Searching using keywords based on few experienced individuals, however, introduces biases and reduce the likelihood of finding research one does not know. One possible solution is text mining which allows automatic determination of relevant search terms based on a large data set (cf. Grames, Stillman, Tingley & Elphick, 2019a). As part of a systematic review on the effect of language-sensitive subject teaching approaches (Vasylyeva, Woerfel, Twente & Höfler, in prep), a text mining method based on co-occurrence networks was used to optimize the search strategy. Using the R package "litsearchr" (Grames, Stillman, Tingley & Elphick, 2019b), terms that occur together and have a specific binding strength were identified in a collection of naively searched literature (2668 documents of a free and controlled search in Scopus and ERIC and 58 in a FIS Bildung). These terms, which are particularly representative of the content of relevant documents, supplement a multiple-stage search term selection process. This presentation presents the application of litsearchr in a German- and English-language SR in language education and discusses problems and benefits of the application.

Authors: Ruszkowski, Jakub; Witkowski, Jacek M; Dębska-Ślizień, Alicja
Title: Key role of citation chasing in the evidence synthesis on the gastrointestinal symptoms prevalence in chronic kidney disease: a case study
Abstract: Even though meta-analyses are mainly performed to establish the effectiveness of various treatments on health and social outcomes, they can also be conducted to improve our understanding of patients' experience of a disease. In the recently published systematic review and meta-analysis on lower gastrointestinal symptoms in patients with chronic kidney disease, we showed that citation chasing (using the Citationchaser app.) of articles introducing symptom questionnaires is an essential step in data collection as a quarter of the papers would not have been found using the standard database search method. Results for each prevalence outcome expressed as single proportions were pooled and visualized using the "meta" package. Using the "altmeta" package, we showed that both a conventional two-step method (with a Freeman–Tukey double arcsine transformation) and generalized linear mixed models (regardless of the choice of link function: logit, probit, cauchit, cloglog) provide relatively similar results. To assess "reporting biases" such as selective non-publication (publication bias) and selective non-reporting of results, we conducted a Peters' regression test, calculated the Luis Furuya-Kanamori index, and generated both funnel and Doi plots using functions from the "meta" and the "metasens" packages. To sum up, our case supports using R packages and shiny apps to conduct a meta-analysis of prevalence.

—

Presentations 3: Searching and record management 2

Authors: Wilson, Emma; Hair Kaitlyn; Macleod Malcolm
Title: The benefits of using R for systematic review reference management
Abstract: Good reference management is essential when conducting systematic reviews or other evidence synthesis research. Many different reference management software programs are available to researchers, including Zotero, Mendeley, EndNote and Papers. However,

reference management software can often struggle to handle large numbers of references, and the lack of version control means that changes made to references may not be recorded or reproducible. Many features of reference management software relevant to systematic reviews – such as importing references from databases searches, storing and organising references, filtering references, and retrieving full-text documents – can be performed using the R programming language. Additionally, changes made to references can be documented in a reproducible way using R scripts or RMarkdown files and GitHub. In this talk, I discuss how R can be used to effectively manage systematic review references, outline the benefits of using R to do so, and show examples from my own systematic review projects.

Authors: Kapp, Claudia; Hausner, Elke; Fujita-Rohwerder, Naomi; Sieben, Wiebke; Lilienthal, Jona; Waffenschmidt, Siw
Title: Text analysis for search strategies: implementing an approach with R
Abstract: As information specialists, our job is to think about how to create and optimize search strategies for systematic reviews. During this presentation, we will introduce the first version of a new R package which implements an updated approach for search strategy development based on Hausner et al., 2012 (1). We will discuss why we chose to create our own package, although we are all novices to programming. Furthermore, we will address what needs we see for new packages and tools for evidence synthesis from the perspective of information specialists.
1. Hausner E, Waffenschmidt S, Kaiser T, Simon M. Routine development of objectively derived search strategies. Syst Rev. 2012;1:19.

Authors: Zhang, Qiyang; Pallath, Akash
Title: Paperfetcher: A tool to automate handsearching and citation searching for systematic reviews
Abstract: Systematic reviews are vital instruments for researchers to understand broad trends in a field and synthesize evidence on the effectiveness of interventions in addressing specific issues. The quality of a systematic review depends critically on having comprehensively surveyed all relevant literature on the review topic. In addition to database searching, handsearching is an important supplementary technique that helps increase the likelihood of identifying all relevant studies in a literature search. Traditional handsearching requires reviewers to manually browse through a curated list of field-specific journals and conference proceedings to find articles relevant to the review topic. This manual process is not only time-consuming, laborious, costly, and error-prone due to human fatigue, but it also lacks replicability due to its cumbersome manual nature. To address these issues, we present a free and open-source Python package and an accompanying web-app, Paperfetcher, to automate the retrieval of article metadata for handsearching. We will also demonstrate how Paperfetcher can be used in R! With Paperfetcher's assistance, researchers can retrieve article metadata from designated journals within a specified time frame in just a few clicks. In addition to handsearching, it also incorporates a beta version of citation searching in both forward and backward directions. Paperfetcher has an easy-to-use interface, which allows researchers to download the metadata of retrieved studies as a list of DOIs or as an RIS file to facilitate seamless import into systematic review screening software. To the best of our knowledge, Paperfetcher is the first tool to automate handsearching with high usability and a multi-disciplinary focus.

Authors: Haddaway, Neal
Title: GSscraper and greylitsearcher - useful but flawed tools for searching for studies in evidence syntheses

Abstract: Web scraping is a useful technique for extracting patterned data when searching for studies in evidence syntheses. It holds promise where search results cannot be exported directly in bulk, and allows data to be integrated into eligibility screening pipelines. Here, I report on two tools (GSscraper and greylitsearcher) build using basic web scraping in R and hosted as Shiny apps. I explain the problems associated with these methods and call for additional support in helping to make these web scraping tools resilient to code fluctuations in the underlying websites (Google and Google Scholar).

—

Presentations 4: Quantitative synthesis 1

Authors: Li, Ranran; Balliet, Daniel; Thielmann, Isabel; De Vries, Reinout
Title: Meta-Analyzing Variation: Do Strong Situations Restrict Variance in Behaviors?

Abstract: As meta-analysis has become a standard approach to synthesizing empirical findings, most meta-analyses have focused exclusively on comparing mean differences between two groups. An overlooked yet as well important effect lies in the variance. An experimental treatment may affect not only the mean but also the variance; investigating variance differences between groups enables us to gain fresh insights from existing datasets. Some metrics for the meta-analytic comparison of variability have been proposed, lnCVR, for instance (Nakagawa et al., 2015; Senior et al., 2020), which has been incorporated in the metafor package (v3.8-1; Viechtbauer, 2010). Here we applied this technique to examine the situational strength hypothesis. The situational strength hypothesis, positing that strong situations restrict behavioral variance, has been treated as a maxim in personality-social psychology and organizational behavior. Prior work has, however, offered inconclusive support for the hypothesis. Here, we provide a systematic evaluation of the hypothesis using a preregistered meta-analysis (k = 128, N = 25,670) to examine the proposed restricted variance effect on prosocial behavior as assessed in economic games modeling social dilemma situations. Strong versus weak situations were represented via multiple theoretically relevant game features. Results showed that strong situations indeed produced restricted behavioral variance when compared to weak situations (overall meta-analytic effect MLMA b_lnCVR = -.26 [-.33, -.18]). Furthermore, subjective perceptions from independent raters (k = 41, n_conditions = 138) corroborated the findings: situations hypothesized and/or perceived as strong (by independent raters) did indeed restrict behavioral variance. This large-scale investigation using economic games thus corroborates the situational strength hypothesis.

Authors: Dolgorukova Antonina, Protsenko Ekaterina, Isaeva Julia, Gagloeva Victoria, Verbitskaya Elena, Berkovich Regina, Sokolov Alexey Y
Title: A meta-analysis of preclinical studies with complex data structure: a practical example of using a multilevel model to account for dependencies

Abstract: The low reliability and reproducibility of preclinical studies findings indicate the need for meta-analytic research allowing not only more accurate estimates, but also the

identification of risks of bias, publication bias, and design features potentially affecting the results. The common challenge of preclinical meta-analyses is the complex data structure implying dependent effect sizes, which, if ignored, can result in misleading statistical inferences. Multilevel modelling and robust variance estimation are the most reliable approaches for handling dependencies, however, they have not yet been widely adopted. Here we demonstrate a practical example of the application of these methods in the meta-analysis of controlled studies testing migraine treatments in the animal model of trigeminovascular nociception (study protocol at PROSPERO: CRD42021276448).
Our systematic search identified 13 studies reporting on 21 experiments, some of which used a shared control group. A three-level model with robust variance estimation was built using the rma.mv() and robust() functions of the metafor package for R. The extent to which methodological features and the reporting of measures to reduce bias explain the observed heterogeneity was assessed in subgroup analyses (meta-regression). To test the robustness of the results, we also examined the presence of outliers and influential cases, followed by sensitivity analysis and estimated potential publication bias. We believe that this work is a helpful example of using the metafor package for multilevel modelling in preclinical meta-analyses and would like to discuss the used methodology and results.

Authors: Harris, Rebecca; Batterham, Marijka; Neale, Elizabeth; Ferreira, Isabel
Title: Pre-eclampsia in pregnancy and offspring blood pressure: a multilevel multivariate meta-analysis of observational studies
Abstract: Background: In studies pertaining to cardiovascular health, systolic and diastolic blood pressure are key outcomes of interest and are both usually reported in primary studies. When conducting meta-analysis, such outcomes cannot be combined in a standard pairwise meta-analysis because they are not independent. An appropriate approach for addressing multiple dependent outcomes in meta-analysis is through multilevel modelling which accounts for the correlation by specifying how each effect size is nested in the included studies. We present a case-study where we used multilevel meta-analysis to analyse multiple outcomes (systolic and diastolic blood pressure) and multiple follow-up measures from cohort studies assessing the impact of pre-eclampsia on offspring blood pressure. Methods and Results: To identify articles, we searched the Medline (via PubMed), CINAHL (via EBSCO) and Embase (via Elsevier) databases from their inception to January 31, 2022. Meta-analysis of 42 effect sizes from 12 studies was conducted using the metafor package in R. When analysing effect sizes adjusted for confounders, offspring exposed to a pre-eclamptic pregnancy had higher systolic (SMD: 0.157; 95%CI: 0.098, 0.216) and diastolic blood pressure (SMD: 0.136; 95% CI: 0.068, 0.203). Compared to univariate pairwise meta-analysis, pooled effects from multilevel multivariate analyses were stronger and precision around DBP was greater. Results from meta-regression tests to compare early and late onset of pre-eclampsia were not statistically significant. Conclusions: This multilevel meta-analysis confirms the positive association between pre-eclampsia and offspring blood pressure after accounting for potential confounders while accounting for the multilevel structure of the data.

Authors: Viechtbauer, Wolfgang
Title: Location-scale models for meta-analysis using the metafor package

Abstract: The purpose of most meta-analyses is to estimate the size of the average effect and/or to examine under what circumstances the effect tends to be higher/lower. However, equally important is the question how much the effect varies across studies. The latter question is focused on the amount of heterogeneity in the effects, which we can estimate under a random-effects model using well-established methods. However, these methods assume that the amount of heterogeneity does not depend on the study characteristics and hence is assumed to be constant (homoscedastic). An extension of the standard random-effects model - the meta-analytic location-scale model - relaxes this assumption and allows researchers to examine under what circumstances the amount of heterogeneity tends to be higher/lower. In this tutorial, I will demonstrate how such location-scale models can be fitted using the metafor package and discuss the potential and limitations of such models.

Authors: Nakagawa, Shinichi; Noble, Daniel; Lagisz Malgorzata; Spake Rebecca; Viechtbauer, Wolfgang; Senior, Alistair
Title: Meta-analyses with missing standard devisations with log response ratio
Abstract: The log response ratio, lnRR, is the most frequently used effect size statistic in ecology. However, missing standard deviations (SDs) are often present in meta-analytic datasets, preventing us from obtaining the sampling variance for lnRR. We propose three new methods to deal with missing SDs. All three methods use the square of the weighted average coefficient of variation CV to obtain sampling variances for lnRR when SDs are missing. Using simulation, we find that using the average CV to estimate the sampling variances for all observations, regardless of missingness, performs best. Surprisingly, even where SDs are missing, this simple method performs better than the conventional analysis with no missing SDs. This is because the conventional method incorporates biased estimates of sampling variances as opposed to less biased sampling variances with the average CV. All future meta-analyses of lnRR could take advantage of our new approach along with the other methods.

Authors: Patzl, Sabine; Diedrich, Jennifer; Pietschnig, Jakob; Lewalter, Doris
Title: Using multiverse and specification curve analyses as an assessment of generality of effects for MASEMs: A meta-analysis on creative potential and self-assessment measures
Abstract: Creativity is not only a characteristic of, e.g., scientific geniuses and artists, but it is the general opinion that everyone has a certain amount of creative potential. A person's creative potential does not necessarily lead to creative achievements, as there is evidence that the relationship might be partially explained through creative self-assessments (=CSA) such as creative self-beliefs. However, the question of the relationship between CSA and the actual creative potential remains unresolved. The main goal of this meta-analysis is to investigate whether two indicators of creative potential (i.e., divergent thinking and intelligence) are associated with CSA. Here, we use a meta-analytical structural equation modeling (=MASEM) approach as proposed by Wilson and colleagues (2016), to model expectable effect size dependencies. Furthermore, we expect to find a substantial amount of heterogeneity in the data due to effects of moderating variables. However, because MASEMs are limited in how many moderators can be included, we will use two approaches to investigate the effect generality of the scrutinized MASEM. First, we will apply subgroup analyses to test if parameter estimates are equal across the different CSA types and age groups. Second, we will apply multiverse and specification curve analyses to all bivariate

relationships. This allows us to investigate the influence of various study design variables and (reasonable) meta-analytical decisions simultaneously and thus contributes to disentangling the causes of inconsistent study results concerning the relationship between creative potential and CSA in the available literature. We show how multiverse and specification curve analyses combined with MASEMs can be used to assess the generality of research synthesis outcomes.

—

Presentations 5: Quantitative synthesis 2

Authors: Heck, Daniel W
Title: metaBMA: Bayesian Model Averaging for Meta-Analysis in R
Abstract: Meta-analysis aims at the aggregation of observed effect sizes from a set of primary studies. Whereas fixed-effect meta-analysis assumes a single, underlying effect size for all studies, random-effects meta-analysis assumes that the true effect size varies across studies. Often, the data may not support one of these assumptions unambiguously, especially when the number of studies under consideration is small. In such a case, selecting one of the two models results in too narrow confidence intervals when assuming fixed-effects but in low statistical power when assuming random-effects. As a remedy, Bayesian model averaging can be used to combine the results of four Bayesian meta-analysis models: (1) fixed-effect null hypothesis, (2) fixed-effect alternative hypothesis, (3) random-effects null hypothesis, and (4) random-effects alternative hypothesis. Based on the posterior probabilities of these models, Bayes factors allow to quantify the evidence for or against the two key questions: "Is the overall effect non-zero?" and "Is there between-study variability in effect size?". Besides considering model uncertainty, Bayesian inference enables researchers to include studies sequentially in order to update a meta-analysis as new studies are added to the literature. The R package metaBMA facilitates the application of Bayesian model-averaging for meta-analysis by providing an accessible interface for computing posterior model probabilities, Bayes factors, and model-averaged effect-size estimates for meta-analysis.

Authors: Braginsky, Mika; Mathur, Maya
Title: Tools for accounting for within-study and across-study biases in meta-analysis
Abstract: Meta-analytic estimates can be systematically biased when the meta-analyzed studies are subject to within-study biases (e.g. confounding) and/or across-study biases (e.g. publication bias and phacking). Existing methods are limited in that they generally analyze only one bias at a time and can make unrealistic statistical assumptions. We present a novel toolkit for addressing these issues by conducting meta-analytic bias correction and sensitivity analysis. Our methods allow taking into account the separate and joint effects of both within-study and across-study biases. These tools allow the researcher to answer questions such as: "For a given severity of the bias(es) in question, how much would the estimate change?" and "How severe would the bias(es) have to be to attenuate the estimate to the null?". The toolkit includes the R packages PublicationBias, phacking, and multibiasmeta, as well as accompanying Shiny apps (which can be accessed at metabias.io). The PublicationBias and phacking packages provide analyses for the corresponding across-study biases, while the multibiasmeta package provides analyses for the joint effects

of publication bias and various within-study biases. We demonstrate how to conduct these analyses both in R code and in Shiny apps.

Authors: Fuenderich, Jens; Frank, Maximilian; Beinhauer, Lukas
Title: MetaPipeX: Data analysis & harmonization for multi-lab replications of experimental designs
Abstract: The number of multi-lab replication studies (e.g. ManyLabs, Registered Replication Reports) in psychology is gradually increasing, with few uniform standards in data preparation or provision. This leads to challenges in both access and re-use of multi-lab replication data. The MetaPipeX framework takes on these challenges, serving as a novel proposal to standardize data structure, analysis code and reporting for experimental data of between groups comparisons in replication projects. It provides users with both structure and tools to synthesize and analyse mulit-lab replication data, select relevant subsets or create helpful graphics such as violin-, forest- and funnel-plots. MetaPipeX consists of three components: A descriptive pipeline for data transformations and analyses, analysis functions that implement the pipeline and a Shiny App utilizing the standardized structure for insights into the data at different aggregation levels. The analysis functions are largely built around meta-analysis of effect sizes (components), utilizing the metafor::rma.mv function (Viechtbauer, 2010). The analysis results consist of replication statistics, meta-analytical model- and heterogeneity-estimates. Additionally the functions provide documented data exports of various agreggation levels. All kinds of data subsets or graphics may be exported for further use. In this tutorial at ESMARConf we will present the framework and show personas ("prototypical users") with different use cases ranging from data analytical tasks to educational purposes. In order to contextualize the framework and its features we will provide a brief summary of the current state of repositories from multi-lab replication projects and discuss potential benefits and limitations of standardization. Using the MetaPipeX framework, we aim to save other researcher countless hours of data manipulation and harmonization, building a foundation for future reproducible multi-lab replication studies.

Authors: Pustejovsky, James E.; Joshi, Megha
Title: Clustered bootstrapping for handling dependent effect sizes in meta-analysis: Exploratory application for publication bias analysis
Abstract: In many fields, quantitative meta-analyses involve dependent effect sizes, which occur when primary studies included in a synthesis contain more than one relevant estimate of the relation between constructs. When using meta-analysis methods to summarize findings or examine moderators, analysts can now apply well-established methods for handling dependent effect sizes. However, very few methods are available for examining publication bias issues when the data also include dependent effect sizes. Furthermore, applying existing tools for publication bias assessment without accounting for effect size dependency can produce misleading conclusions (e.g., too-narrow confidence intervals, hypothesis tests with inflated Type I error). In this presentation, we explore a potential solution: clustered bootstrapping, a general-purpose technique for quantifying uncertainty in data with clustered structures, which can be combined with many existing analytic models. We demonstrate how to implement the clustered bootstrap in combination with existing publication bias assessment techniques like selection models, PET-PEESE,

trim-and-fill, or kinked meta-regression. After providing a brief introduction to the theory of bootstrapping, we will develop and demonstrate example code using existing R packages, including `boot` and `metafor`. Time permitting, we will also share findings from ongoing methodological studies on the performance of clustered bootstrap selection models.

Authors: Jones, Matt Lloyd
Title: Three challenges from a recent meta-analysis and how I tried to deal with them
Abstract: Each meta-analysis presents its own set of unique challenges that the meta-analyst must seek ways of dealing with – especially as meta-analysis is increasingly being applied in non-medical fields where study designs and reporting standards are often more diverse. Here, Matt will present some reflections on his experience conducting a meta-analysis of veterinary microbiology studies of the antibiotic use-resistance relationship in beef cattle. In this field, issues such as multiple publications related to the same study, unit of analysis problems, and the use of proportional measures of the outcome present particular challenges. Matt will share his reflections on dealing with these challenges using R in the hope of stimulating discussion around these issues that may help others or himself better deal with them in the future!

—

Presentations 6: Data visualisation and communication

Authors: Evrenoglou, Theodoros; Boutron, Isabelle; Seitidis, Georgios; Ghosn, Lina; Chaimani, Anna
Title: metaCOVID: A web-application for living meta-analyses of Covid-19 trials
Abstract: COVID-NMA is an international initiative that performs 'living' evidence synthesis for all treatments and vaccines used against Covid-19. Through its platform COVID-NMA provides access to the most-up-to date findings regarding more than 300 treatment comparisons and more than 20 vaccines. The initiative has received recognition by important organizations such as WHO and Cochrane while many guideline developers have declared their engagement to the outputs of the platform. However, apart from real time access to the data, stakeholders also need to investigate the data and the impact of different characteristics on the results as well as to produce their preferred evidence summaries. To assist them, we developed and made freely available the metaCOVID application. This web-application allows the end-users of the COVID-NMA platform and other external researchers to directly use the most up-to-date database and perform meta-analyses tailored to their needs in a user-friendly environment. The users can interact with the data and customize their analysis by clicking one or more of the buttons which are available in the user interface. Based on their selection the default analysis can be modified in many different ways: (a) type of meta-analysis model (b) method for heterogeneity estimation (c) subgroup analysis criteria (d) exclusion of pre-prints from the analysis (e) exclusion of studies according their risk of bias status (f) Hartung-Knapp adjustment for the confidence intervals. Analyses are performed using the R-package metafor and the results are presented through downloadable forest plots. Those forest plots are enriched with several study characteristics as well as a risk of bias assessment for each study. In summary, metaCOVID offers open access to the most-up-to date database of Covid-19 trials for researchers, clinicians, or

guideline developers interested to perform amendable meta-analyses and explore the impact of certain characteristics on the results.

Authors: Lukas Röseler, Lukas Wallrich, Brad J. Bushman
Title: Creating interactive ShinyApps for Meta-Analyses with metaUI
Abstract: Meta-analyses are based on rich datasets that can be analyzed in numerous ways, and it is unlikely that authors and readers will always agree on the "best ways" to analyze the data. Whether it comes to the choice of model (e.g., random versus fixed effects), the methods for assessing or adjusting for publication bias (e.g., z-curve, p-curve, robust Bayesian meta-analysis), or the moderators to be included in a meta-regression, disagreements are likely to arise. This can lead to the inclusion of lengthy robustness checks and alternative analyses that are time-consuming and difficult to digest. Here we present metaUI, a new R package that supports researchers in creating an interactive web app that allows readers (and reviewers) to explore meta-analytic datasets in a variety of different ways. Apart from allowing readers (and reviewers) to assess the robustness and trustworthiness of results more comprehensively, metaUI allows others to assess the results that are most relevant to them, such as by filtering the dataset to focus on a specific group of participants, region, outcome variable, or research method. With the opportunity for users to download the dataset used and to upload alternatives, it will also facilitate the updating of meta-analyses. To date, some researchers have created similar web apps for their meta-analyses that have been well received, yet they require substantial time investment and advanced coding skills to use. With metaUI, researchers can get a working app by simply uploading their dataset and tagging key variables – while they still have the flexibility to tailor the display in line with their interests and requirements. In this session, we demonstrate the use of the package by creating interactive apps for illustrative datasets from the psymetadata package and collect initial feedback for further development."

Noble, Daniel; Nakagawa, Shinichi; Senior, Alistair; Malgorzata, Lagisz; O'Dea, Rose; Rutkowski, Joanna; Yang, Yefeng   Making orchaRd plots for meta-analysis   Classic forest plots in meta-analyses are often of limited use when there are hundreds of effect sizes. We suggest that a new plot, called an orchard plot, is more useful across a broad array of meta-analytic research because it not only provides aggregated meta-analytic means along with 95% confidence intervals within sub-groups, but it also visualises the raw effect size data (scaled by their precision) and 95% prediction intervals. The 95% prediction intervals allow readers to understand the range of effect sizes expected from future studies and it's the most ideal measure of heterogeneity in a meta-analysis. We overview the functionality of our new R package, orchaRd, to show how it can be used to make orchard plots.

Authors: Georgios, Seitidis; Sofia, Tsokani; Christos, Christogiannis; Katerina-Maria, Kontouli; Alexandros, Fyraridis; Stavros, Nikolakopoulos; Areti Angeliki, Veroniki; Dimitris, Mavridis
Title: Graphical tools for visualizing the results of network meta-analysis of multicomponent interventions
Abstract: Network meta-analysis (NMA) is an established method for assessing the comparative efficacy and safety of competing interventions. It is often the case that we deal with interventions that consist of multiple, possibly interacting, components. Examples of interventions' components include characteristics of the intervention, mode (face-to-face,

remotely etc.), location (hospital, home etc.), provider (physician, nurse etc.), time of communication (synchronous, asynchronous etc.) and other context related components. Networks of multicomponent interventions are typically sparse and classical NMA inference is not straightforward and prone to confounding. Ideally, we would like to disentangle the effect of each component to find out what works (or does not work). To this aim, we propose novel ways of visualizing the NMA results, describe their use, and illustrate their application in real-life examples. We developed an R package viscomp to produce all the suggested figures.

Authors: Chris C Pritchard, Neal R Haddaway
Title: Significant updates to the PRISMA2020 package supporting use as an API
Abstract: tba