

Exploratory Analyses for Missing Data in Meta-Analyses

Introduction

Systematic reviews of substance abuse research hold great promise for examining what makes potential interventions effective (White *et al.*, 2010; Tanner-Smith, Wilson, & Lipsey, 2013; Tanner-Smith *et al.*, 2016; Newbury-Birch *et al.*, 2018; Ramsey *et al.*, 2019; Yuvaraj *et al.*, 2019). Methodological tools such as meta-regression can formally test relationships between and intervention’s impact and how or on whom it is implemented (Hedges & Olkin, 1985; Cooper, Hedges, & Valentine, 2019; Tipton, Pustejovsky, & Ahmadi, 2019a). However, such tools must contend with the real-world difficulties of modern research syntheses, including the fact that it is often impossible to extract all relevant information from the literature to conduct such analyses.

The fact that not every study reports the information required to run a meta-regression means that many meta-analyses face missing data problems (Pigott, 2019). Issues with missing data are not new. There is a vast literature on methods for handling missing data in primary studies, as well as work on related issues in meta-analysis (Rubin, 1976; Pigott, 2001a, 2019; Schafer & Graham, 2002; Little & Rubin, 2002; Graham, 2009; van Buuren, 2018). This literature highlights the ways that missingness can bias an analysis, examines conditions under which these biases can be corrected, and proposes various statistical procedures to adjust for bias or accurately compute uncertainty.

Diagnosing missing data issues remains an important aspect of any analysis and presentation of incomplete data (i.e., data with missing values). Understanding which and how much data is missing, and how problematic that is, is crucial in determining how to proceed in a meta-analysis and how to contextualize the results. A key assumption of many analysis methods for incomplete data is that the analyst has some idea about why data is missing. While much of the literature has focused on the implications of that assumption, considerably less attention is paid to approaches to examining it in a dataset (Tierney & Cook, 2018).

Recent research has suggested analysts can better understand missingness in their data through exploratory analyses, including visual and numerical summaries akin to classical exploratory data analyses (Tukey, 1962; Buja, Cook, & Swayne, 1996; Cheng, Cook, & Hofmann, 2015). These explorations, which occur before conducting the formal meta-analysis, can shed greater light on key issues relevant to missingness. Tools for doing so are only now emerging, but these tools have yet to gain broader traction in quantitative disciplines (Tierney, 2017; Tierney & Cook, 2018). Nor has this approach seemingly made its way into meta-analysis, where missing data is a common problem.

This tutorial demonstrates tools for exploring and diagnosing missing data problems in meta-analysis. The following section clarifies the types of missing data for which these tools are appropriate. We then describe principles of missing data that can guide exploratory analyses. Finally, we show some of these tools on data from a meta-analysis on substance abuse interventions for adolescents. Additional examples and executable code are available as part of the supplementary material.

Missing Data in a Meta-Analysis

Because a meta-analysis involves an ensemble of *effects* (i.e., intervention impacts) reported by primary studies, *missing data* or *missingness* in meta-analysis could refer to at least three different scenarios. For instance, data could be missing on individual participants within studies, including their outcomes in the study or other characteristics (e.g., their age, race, prior substance use) (e.g., Higgins, White, & Wood, 2008). Missingness could also refer to information that could not be extracted from a completed study by a meta-analyst (Pigott, 2001a). This may occur if a study fails to report enough detail for analysts to back out effect estimates, standard errors, or study- and effect-level characteristics. Finally, entire studies or effects may be missing from a meta-analytic dataset. This might occur if effects (or entire studies) are

not reported or published (Rosenthal, 1979). There is empirical evidence that statistically significant results are more likely to be published and hence wind up in a meta-analysis, which can induce *publication bias*, a well-known problem in the field (Hedges, 1984; Rothstein, Sutton, & Borenstein, 2005). The studies or effects that are not reported, and thus are not included in a meta-analysis, can be seen as missing data.

Precisely how to examine, diagnose, and adjust for missing data will be different depending on what scenario we mean when we say “missing data.” For instance, meta-analysts have used *funnel plots* to assess whether their systematic review is missing studies or effects due to publication bias (Light & Pillemer, 1984; Egger *et al.*, 1997).

Our focus will be on the second scenario, where information cannot be extracted from some studies. This is a common problem in meta-analysis and one that can limit the accuracy of any statistical inferences (Pigott, 2019; Tipton, Pustejovsky, & Ahmadi, 2019b). One can think of this type of missingness in terms of a dataset or table. Assume we have data on k effect estimates and p variables (including the estimate itself). This can be summarized and stored in a $k \times p$ table where rows correspond to effect estimates and columns correspond to variables concerning those estimates. One column would contain the effect estimates themselves, and another would contain the standard error or estimation error variance of those estimates. The remaining $p - 2$ columns could contain effect- or study-level covariates, including summary demographics (e.g., the percent of a study’s sample that were minorities), treatment type (e.g., behavioral therapy versus pharmacological interventions), or dosage/duration of an intervention. *Missing data* in this context would refer to individual cells in this dataset that are missing values.

Data

A prime example of this type of missingness can be seen in data from Tanner-Smith *et al.* (2016), who examined the impacts of substance abuse interventions for adolescents on subsequent substance use. These data were extracted from 61 randomized trials and quasi-experiments, and include $k = 95$ different effect size estimates. These data will be used to illustrate key concepts of missingness and some useful tools to exploring missingness in this tutorial.

Tanner-Smith *et al.* identified a range of intervention types and venues, and those interventions have been studied on individuals who use different substances and who differ in a variety of ways. Some interventions in their data focus on cognitive behavioral therapy (CBT), family therapy, or pharmacological therapy. Some interventions are in-patient, and others are out-patient. Individuals in studies might present using marijuana, alcohol, or opioids, and they may come from wealthy families or poor families. Finally, some effects reported in studies contrasted a given intervention with some control condition, while others contrasted two alternative interventions or implementations.

To explore relevant relationships, Tanner-Smith *et al.* extracted a considerable amount of information from the studies they found. Their raw data included some $p = 46$ variables per study. In addition to estimated effects and their standard errors, they documented the types of interventions being contrasted, as well as their intensity and context. This included where interventions occurred, and how much time subjects spent in the intervention. For instance, if a study contrasted two interventions, Tanner-Smith *et al.* documented how many hours per week subjects in each intervention (referred to in this article as *groups*) spent in receiving treatment. They also documented the demographics of subjects in the studies, such as the percentage of subjects who were minorities, as well as the substances that subjects reported using.

Tanner-Smith *et al.* then fit a series of meta-regression models to their data in order to examine how treatment impacts varied according to the type of therapies and individuals studied. They found that assertive continuing care (ACC), behavioral therapy, CBT, motivational enhancement therapy (MET), and family therapy tended to be more effective than generic “practice as usual” interventions that often involved referrals to community services. However, they did not find strong relationships between the characteristics of adolescents in the studies and the effectiveness of interventions (net of intervention type).

A complicating factor in conducting these analyses was that some of the data were missing. Not every study reported the requisite information for extracting covariates for every effect size. For instance, not all studies reported how many hours per week subjects spend in therapy or the racial or socioeconomic makeup of their subject pool. As a result, not all effect estimates had information about the types of individuals in the study or the intensity of the interventions. It was often the case that one or two of the

fields in their dataset were missing for any given effect estimate. Thus, when it came time to run meta-regressions, Tanner et al. were faced with a decision about how to address the information that was missing.

Principles of Missing Data

Tanner-Smith et al. ultimately opted for a sophisticated statistical procedure called the expectation-maximization (EM) algorithm to estimate their models, which has been an important tool for analyzing data with missing values (Dempster, Laird, & Rubin, 1977; Graham, Cumsille, & Elek-Fisk, 2003). The EM algorithm has also been studied as a useful approach to estimation when missing covariates in a statistical model, which was primarily the issue facing Tanner-Smith et al. (Ibrahim, 1990; Ibrahim, Lipsitz, & Chen, 1999).

However, that was not their only option. A common approach in meta-analysis is a *complete-case* analysis that excludes effects for which any of the relevant covariates in the meta-regression model are missing (Pigott, 2001a, 2001b, 2019; Tipton *et al.*, 2019b). An alternative to complete-case analysis that has gained broad use in various fields is to impute (i.e., fill in) missing values. Pigott (2019) discusses the use of single-value imputations in meta-analysis. A more common approach to missingness in primary studies is to use multiple imputations for a missing value, which can more accurately reflect the uncertainty introduced by filling in unknown values. Often imputations are based on predictive models that can better inform what values we might have observed had a given field not been missing. These predictive models typically leverage information from other variables in the data.

Analyses involving incomplete data will be impacted by which variables are missing in a dataset, how frequently they are missing, as well as relationships between variables. Some understanding of these issues in a dataset will guide decisions about which analytic approach may be appropriate. This section provides an overview of principles of missing data that apply to meta-analysis, and lists some potential statistical approaches to handling missing data. These principles can be used to guide exploratory analyses of missingness.

Notation

As noted above, we can describe the datasets commonly used in meta-analysis as tables with k rows and p columns. Each row corresponds to an effect size, and each column corresponds to some variable related to that effect. These are the types of tables used in most standard meta-analysis software, including Comprehensive Meta-Analysis, the `metafor` library, or OpenMetaAnalyst (Viechtbauer, 2010; Borenstein *et al.*, 2012; Trikalinos, 2012). The structure of these tables is shown in the matrix \mathbf{D} below. In the matrix, T_i denote the effect size estimates and σ_i are their standard errors. The X_{ij} refer to additional variables collected that pertain to a given effect size and that might be used in an analysis.

$$\mathbf{D} = \begin{bmatrix} T_1 & \sigma_1 & X_{11} & \dots & X_{1,p-2} \\ T_2 & \sigma_2 & X_{21} & \dots & \mathbf{NA} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ T_k & \sigma_k & \mathbf{NA} & \dots & X_{k,p-2} \end{bmatrix} \quad \mathbf{R} = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 1 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 0 & \dots & 1 \end{bmatrix}$$

Missing values in the matrix are denoted as \mathbf{NA} . Most missing data literature augments the traditional dataset \mathbf{D} above with a matrix \mathbf{R} of response indicators R_{ij} . These indicators take a value of $R_{ij} = 1$ if entry i, j in the dataset \mathbf{D} is observed, and $R_{ij} = 0$ if it is missing. For instance, $R_{21} = 1$ because the first element of row 2 (T_2) is observed, but $R_{2p} = 0$ because the last element of row 2 is missing (\mathbf{NA}).

In the dataset \mathbf{D} above, there are both *observed* and *unobserved* data. The observed data comprises all of the entries in the table above for which $R_{ij} = 1$. For instance, the observed data would include the effect estimates T_i and standard errors σ_i . The unobserved data comprises all of the entries for which $R_{ij} = 0$, including $X_{2,p-2}$ and $X_{k,1}$. Note that the dataset \mathbf{D} , and hence the incomplete data on which we conduct an analysis, is simply the union of the observed and unobserved data.

One issue with missingness involves how much data is missing. This could refer to several different

quantities. For instance, we could be interested in the total fraction of cells with missing values:

$$\sum_{i=1}^k \sum_{j=1}^p \frac{1 - R_{ij}}{kp}$$

We might also be interested in what percentage of observations are missing a given variable (i.e., how much of each column is missing):

$$\sum_{i=1}^k \frac{1 - R_{ij}}{k}$$

While such quantities are somewhat intuitive and can be used in exploratory analyses of missingness, they are not the only numerical summary of missingness. In a meta-analysis, parameters are often estimated by weighting effect size estimates in a way that those estimates that are most precisely estimated (i.e., that have the smallest standard errors) receive the most weight (Borenstein, 2009; Cooper *et al.*, 2019). The precision with which we estimate important quantities in a meta-analysis, including meta-regression coefficients, will therefore depend on the precision of the studies included in an analysis. That is, we will have better estimates of a meta-regression model if the precision of each effect $1/\sigma_i^2$ is large (Hedges & Olkin, 1985). Missing a variable for an effect estimate with a large standard error (and thus low precision) can potentially be less detrimental than missing the same variable for an effect estimate with a small standard error.

Thus, when standard errors of effect sizes are fully observed, an alternative way to quantify the extent of missingness is with a precision-weighted percentage:

$$\sum_{i=1}^k \frac{(1 - R_{ij})/\sigma_i^2}{\sum_{i=1}^k 1/\sigma_i^2}$$

This describes the percentage of the total precision of effects for which a covariate is missing. If this is greater than the raw percentage, that would indicate that a missingness problem may be more acute because a variable is missing from larger studies.

Missingness Mechanisms

A key assumption that underpins analyses of incomplete data involves why those values are missing, typically referred to as the missingness *mechanism*. Rubin (1976) classified three different possible types of mechanisms. These mechanisms relate the probability that a value is missing to the observed and unobserved data; that is, they relate \mathbb{R} to the observed and unobserved data.

Rubin noted that data could be missing completely at random (MCAR), which means that the probability that a given value is missing is independent of all of the observed or unobserved data. This can be expressed as $P[R|\text{observed data, unobserved data}] = P[R]$, and means the probability that a given value is missing is unrelated to anything observed or unobserved.

Values could be missing at random (MAR), which occurs if the probability that a value is missing depends only on the observed data. This can be expressed as $P[R|\text{observed data, unobserved data}] = P[R|\text{observed data}]$. Note that this differs from MCAR in that missingness might be related to observed values. For instance, suppose studies with larger standard errors are less likely to report the racial composition of their samples. Then, assuming the standard errors are observed, this could be consistent with MAR. It would violate an assumption of MCAR, because missingness is related to an observed value: the standard error of an effect estimate.

Finally, data are said to be missing not at random (MNAR) if the probability that a value is missing depends on unobserved data in some way. This differs from MAR in that missingness depends on unobserved, rather than just observed data. As an example, suppose that studies with larger standard errors and a greater proportion of minorities are less likely to report the racial composition of their samples. Then missingness of racial categories in the data would depend on an observed value (the standard error), but also the racial composition that could itself be missing.

It is worth noting that various researchers have proposed statistical tests of some of these assumptions. For instance, Little (1988) describes a test for MCAR, while other authors have posited tests for

testing whether data are MAR versus MNAR (Baker, Rosenberger, & Dersimonian, 1992; Diggle & Kenward, 1994; Molenberghs, Kenward, & Lesaffre, 1997; Troxel, Harrington, & Lipsitz, 2002). However these tests may not be appropriate for data typically used in a meta-regression. While Little’s test assumes that all variables are continuous and normally distributed, meta-regressions frequently involve categorical covariates. Much of the literature on MAR tests compares specific models for dropout in longitudinal studies, which is almost never an issue for the meta-analyst (Molenberghs *et al.*, 2008; Rhoads, 2012).

Missingness Patterns

In addition to the mechanism, it is often useful to understand which variables are missing together from the same rows. For instance, some rows in the Tanner-Smith *et al.* data are missing the hours of therapy per week for one of the groups, while other rows are missing the hours of therapy per week *and* the percentage of study participants who were minorities. In other words, different rows in the data exhibit different *missingness patterns*. Missingness patterns can be thought of examining relationships within the matrix \mathbf{R} .

Understanding these patterns can give some insight into missingness mechanisms, but it can also help identify variables that might be more or less useful in dealing with issues that arise from missingness (van Buuren, 2018). For instance, it will be difficult to impute values for a variable if it is frequently missing alongside several other variables in the data.

Missing Data Analysis Methods

There has been a large amount of research into methods for analyzing incomplete data (see Graham, 2012; Little & Rubin, 2002; van Buuren, 2018). Pigott (2019) provides a comprehensive overview of methods for handling missing data in a meta-analysis, including complete-case analyses (i.e., only analyzing effects for which all variables are observed), available-case analyses (often called “shifting units of analysis” in meta-analysis), single and multiple value imputation (i.e., filling in missingness with one or more values), and the EM algorithm.

Methods for analyzing incomplete data rely on assumptions about the missingness mechanism. Complete-case analyses may be appropriate assuming MCAR, but not necessarily if data are MAR (Pigott, 2001b; Little & Rubin, 2002). More sophisticated approaches, such as multiple imputation or the EM algorithm, are typically implemented in software in a way that assumes data are MAR but not MNAR (Schafer & Graham, 2002; Graham, 2009; van Buuren, 2018).

Exploratory Analyses

The tools discussed in the remainder of this article facilitate exploratory analyses of missingness in a meta-analytic dataset. Exploratory missingness analyses (EMA) are difficult for several reasons. First, it will often be difficult to draw very strong inferences about missingness mechanisms based on exploratory analyses. Even proposed tests for missingness mechanisms can have misleading results (Molenberghs *et al.*, 2008; Rhoads, 2012; Seaman *et al.*, 2013). Instead, EMA can provide support for or help generate theories that explain missingness in ways that are consistent with the mechanisms described above. Some of these theories may require consultation with data curators and other individuals who extracted information from the studies reviewed.

Second, there is no single visualization or set of metrics guaranteed to provide a complete picture of missingness for all datasets. A plot that is tremendously useful for one dataset may be of less interest in others. Any EMA must rely on knowledge of how data were collected and extracted, and can help leverage that knowledge to examine and propose theories about missingness.

Third, EMA differs from traditional exploratory data analyses because the focus is on the unobserved data. Many software tools, including most graphics software actually deletes observations with missing values, which would eliminate the information EMA seek to understand (Tierney, 2017). Thus, an emerging suite of tools that differ from traditional exploratory data analyses are required (Tierney & Cook, 2018). Moreover, as with the case with much general-use statistical software, adaptations of these tools may be necessary to tailor EMA for meta-analytic datasets.

In the following sections, we show and discuss different types of visualizations and numerical summaries relevant to EMA. These are rooted in an approach to understand the scope and correlates of missingness in Tanner-Smith’s et al. data on substance abuse interventions for adolescents. However, they are not exhaustive, and so as part of the supplementary material to this tutorial, we have included a vignette that presents and describes alternative visualizations and numerical summaries of missingness. Both the demonstration presented in this article and the supplementary vignette are implemented in the **R** software language and draw heavily on the **visdat** and **naniar** libraries with some custom extensions developed specifically for meta-analysis (Tierney, 2017; Tierney & Cook, 2018). Executable code is included with the supplementary materials.

Aggregation Plots

Aggregation plots can be a useful starting point when exploring missingness in a meta-analytic dataset. They visualize the entire dataset as a “heatmap” that indicates which values are missing from which rows. Two examples of aggregation plots are given in Figure 1 below. These plots are laid out exactly like the data: the columns correspond to variables in the data, and rows correspond to effect sizes. Figure 1A colors each column according to the type of variable (e.g., numeric, categorical, etc.) they represent, and gray spots indicate that a value in a given row and column is missing. Figure 1B does not specify the type of variables in the data, but instead shows the total number of cells that contain missing values. That is, out of all of the $k \times p$ cells in the data, 6.2% are missing values. **[NOTE TO KARINA: Maybe we choose just one of these to present, or remove some unnecessary columns like pk_es or studyid or __referencesummary?]**

Aggregation plots provide a high-level picture of missingness in a dataset. They can indicate which columns are complete, such as the columns corresponding to the effect size estimates and standard errors (**es_g** and **se_g**: left side of the **numeric** columns in plot A, left-center in plot B). They also show which columns or groups of columns present plenty of missingness. In particular, Figure 1 illustrates that there are groups of variables regarding one of the treatment arms (group 2) in studies that are missing fairly frequently. In all, 74% of the effect sizes were missing values for at least one variable (i.e., in at least one column).

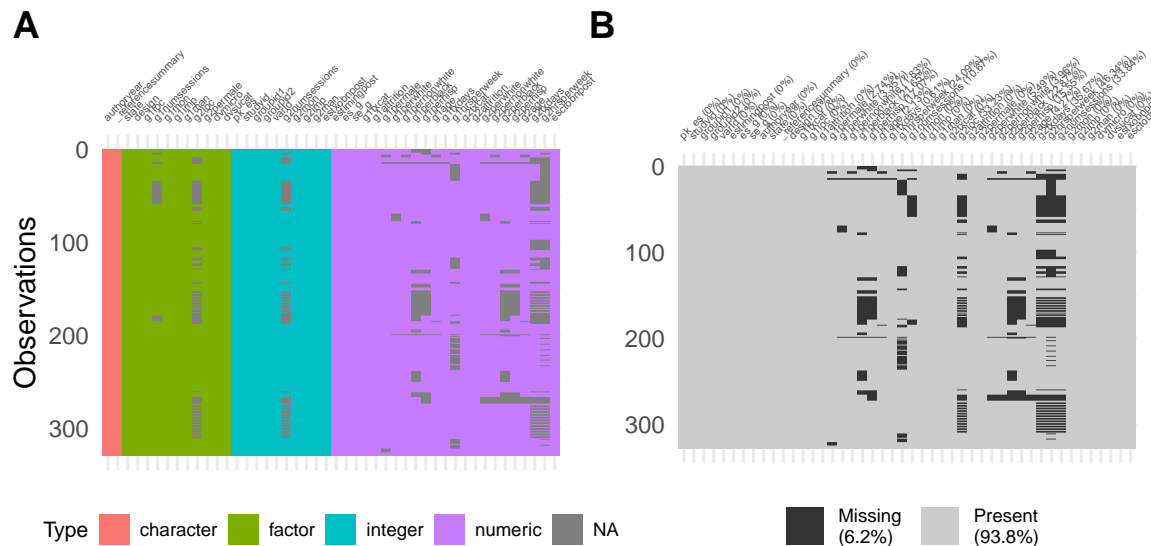


Figure 1: *These plots indicate the severity of missingness in the adolescent substance abuse intervention data. Each row in the plot corresponds to a row in the data, and each column corresponds to a variable collected in the data. Missing cells in the data are indicated by a gray dash in plot (A) and a black dash in plot (B). Plot (A) additionally colors columns according to the type of variables they store. Plot (B) indicates that 6.2% of the total dataset (i.e., 6.2% of all cells in the table) are missing values.*

Univariate Explorations

While aggregation plots can provide a good overview, we typically want more detailed information about how many observations are missing a given variable. Variable missing plots display the overall missingness in each column of a dataset and present the results in order of which column has the most missingness.

Figure 2 presents two approaches to making these plots. Figure 2A displays the raw count missing values in each column, while Figure 2B displays the percent of each column missing. For example, Figure 2 suggests that the hours per week that the contrast group spent in their assigned condition (`g2hrsperweek`) is missing for almost half of the effects in the data. [NOTE TO KARINA: Thoughts on dropping the raw count plot? They’re visually redundant, but I’d be open to keeping both.]

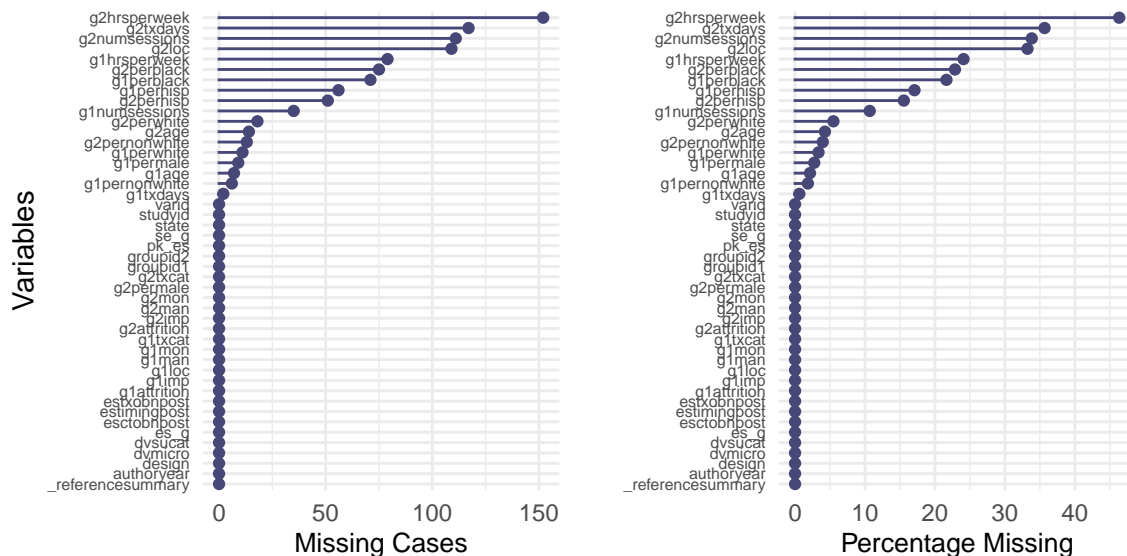


Figure 2: These plots summary missingness in variables, ordered by missingness, in the adolescent substance abuse intervention data. Plot (B) indicates that there are 10 variables with at least 10% of missing cases. This kind of visualization becomes relevant when deciding which variable to include in the analysis.

From variable missing plots, it is often easy to identify variables that might be driving any missing data problems, and they can quantify the extent to which a given column has missing values on the scale of raw percentages. Previously, we argued that precision-weighted percentages may be more informative in describing the extent of missingness in a meta-analysis. Raw percentages and precision-weighted percentages are presented in Table 1. We would discourage interpreting the magnitude of differences between the raw and weighted percentages, however comparing those columns can identify variables missing from effects estimated more precisely. For example, the raw percentage column would suggest missingness in the hours per week that one group in a study spent in treatment (`g1hrsperweek`) may be less of a problem (24% missing). But the weighted percentage indicates that the effects for which it is missing make up nearly 37% of the total precision of effect estimates.

Table 1: This table displays the total number, percentage, and precision-weighted percentage of effect sizes that are missing a given variable.

# Missing	% Missing	Wt. % Missing
152	46.3	45.2
117	35.7	32.6
111	33.8	30.2
109	33.2	30.0

# Missing	% Missing	Wt. % Missing
79	24.1	36.8
75	22.9	17.5
71	21.6	16.1
56	17.1	12.7
51	15.5	12.4
35	10.7	6.9
18	5.5	4.6
14	4.3	3.0
13	4.0	2.3
11	3.4	3.1
9	2.7	1.3
7	2.1	1.5
6	1.8	0.8
2	0.6	0.1

Exploring Patterns of Missingness

When diagnosing potential missing data problems, it also matters which variables are missing together. Certain patterns of missingness can be indicative of the missingness mechanism, and patterns can also point to potential issues for analytic methods with incomplete data, such as multiple imputation, available case analysis, or the EM algorithm.

One approach to visualizing missingness patterns is with an *upset plot* (Conway, Lex, & Gehlenborg, 2017). Upset plots indicate the frequency with which different combinations of variables are missing together. Figure 3 exemplifies an upset plot. The bottom of Figure 3 presents different variables, how many rows each of those variables is missing from, and then describes various missingness patterns. For example, the first missingness pattern corresponds to rows where information about the location (`g2loc_NA`), number of sessions (`g2numsessions_NA`), number of days (`g2txdays_NA`), and hours per week (`g2hrsperweek_NA`) that one of the groups in the study (referred to in the data as “group 2” or `g2` in the plot) spent in treatment are all missing. The top part of Figure 3 describes how frequently those patterns occur in the data. The first pattern, in which much of the information about the treatment implementation is missing for one group in a study, occurs for 38 effects. It is worth noting that variables on the duration, frequency, and timing of treatment in a study are often missing together. Most of the patterns involve combinations of such variables for the same group or for different groups in a study.

Relating Missingness to Observed Values

Explorations of missingness can also examine whether missingness in one variable is related to observed values in another variable. This can highlight potential biases in the observed data. It can also clarify whether data appear to meet the MCAR assumption. The idea behind this is that MCAR assumes that the probability that a value is missing is independent of both observed and unobserved values. However, if missingness in a column in the data is correlated with values in another column, that would be an indication that data are not MCAR. A similar logic underpins the test for MCAR proposed by (Little, 1988).

Since covariates in a meta-regression are often categorical, visualizations like the heatmap in Figure 4 can be useful. Figure 4 demonstrates how missingness in various columns is related to the type of care received in one of the study’s treatment arms. The figure is grouped into tiles or cells. The columns of the figure correspond to the different types of care provided shown along the bottom of the graph: inpatient, outpatient, and continuing care, and some of those values are themselves missing (NA). The rows of the figure correspond to other variables in the dataset. Each cell is shaded according to how frequently those variables are missing for each level of care provided. For example, the percent of subjects who are black or hispanic (`g2perblack` and `g2perhisp`) are much more frequently missing from effects with inpatient treatments than outpatient or continuing care interventions.

Two crucial variables that are in nearly all modern meta-analyses are the effect size estimates and

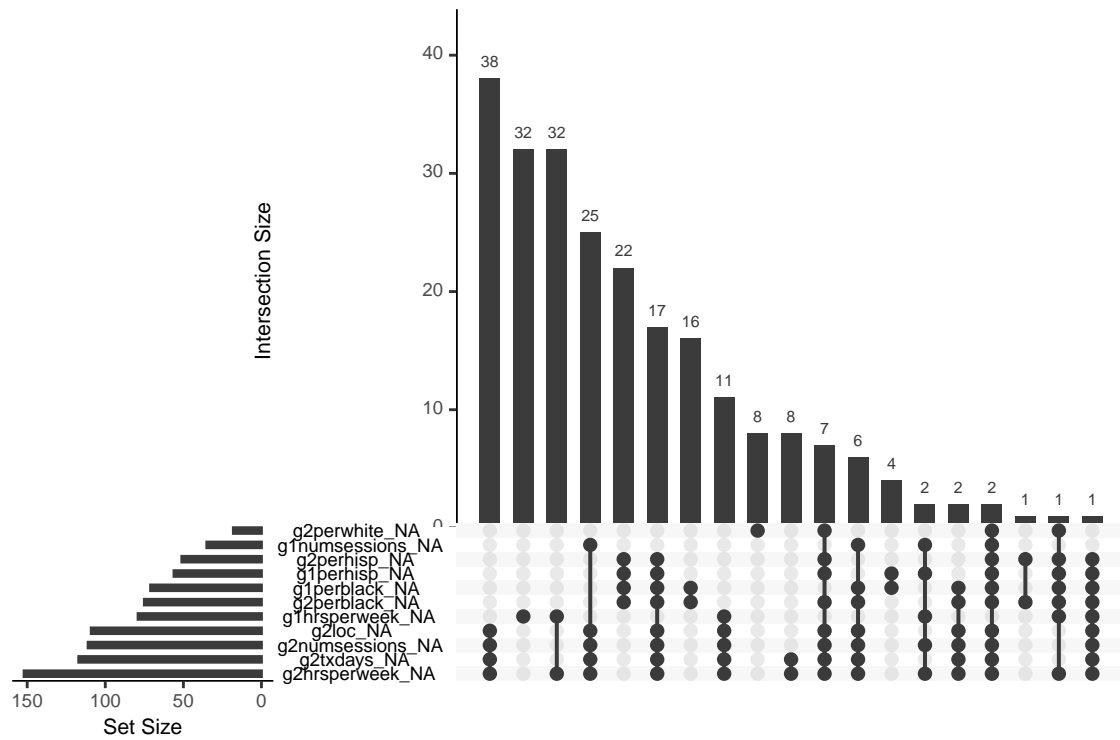


Figure 3: *This plot details those variables that are missing together. For instance, there are a large number of cases where group 2 level of care, number of sessions, treatment contact (hours per week) and duration of treatment (days) are missing together. This simple exploration provides valuable information for imputation.*

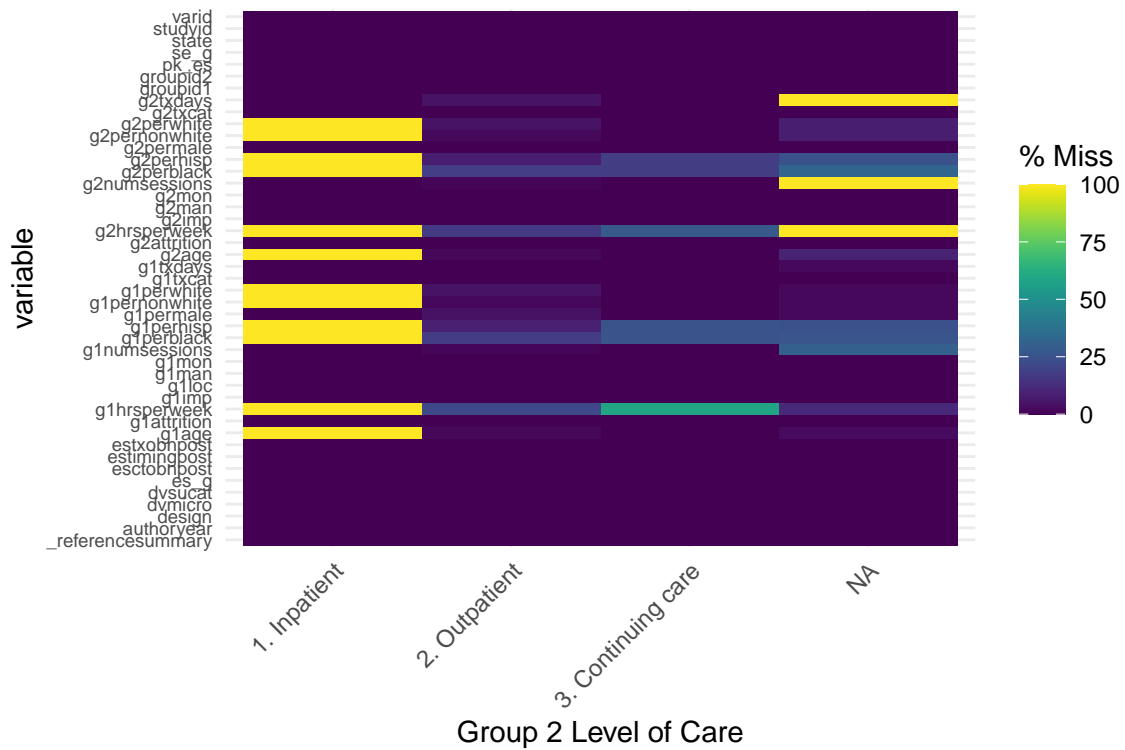


Figure 4: This plot highlights the number of missings in each column, broken down by a factor variable, in this case the level of care for group 2. The inpatient category has 100% of missing values in at least 12 different variables, suggesting that this category could impose a problem when fitting a regression model.

their estimation error variances or standard errors σ_i . If missingness in a covariate is more likely to occur with smaller or larger effect estimates or smaller or larger standard errors, this is likely to impact the analytic results. Therefore, it will typically be a good idea to explore such relationships. Because both effect estimates and standard errors are continuous, density plots may be useful tools.

Density plots present the distribution of effect estimates and standard errors among effects for which a covariate is missing versus when the covariate is observed. Figure 5 displays several density plot pairs for different variables. For each pair of plots we see the distribution of effect estimates (left) and standard errors (right), each colored according to whether a given covariate is missing.

Figure 5A, indicates that effect estimates for which hours of treatment received per week is missing tend to be a little smaller and have smaller standard errors on average than those for which the hours per week were reported. Contrast that with 5B, which shows that effect estimates for which the duration of the treatment in days is missing tend to be larger and have and have larger standard errors than effects for which duration of treatment is reported. Bot plots A and B strongly suggest that missingness of information about treatment dosage will be related to the size of effects found and how precisely those effects were estimated.

Discussion

Missing data is and will continue to be an issue with most meta-analyses, and that can affect what we can learn about substance abuse interventions from research syntheses. While there are various potential approaches to handling missing data in meta-analysis, most of those approaches assume that the missingness mechanism is known to the analyst. That is not always the case, and so this article argued that an exploratory analysis of missingness might help analysts better understand and diagnose their missing data problems. It also outlined and demonstrated some tools that can support exploratory analyses into the scale and correlates of missingness in a meta-analytic dataset. These tools proved to be useful as a first step to understanding why data is missing.

These tools were applied to data on a large meta-analysis conducted by Tanner-Smith *et al.* (2016) on substance abuse interventions for adolescents. We found 74% of the effect sizes were missing at least one of their corresponding covariates. This was driven by some variables that were missing frequently (e.g., Group 2 hours of treatment per week **g2hrsperweek**). Our analysis also revealed that missingness in some variables may be more severe than was obvious from first glance (e.g., Group 1 hours of treatment per week **g1hrsperweek**). Variables quantifying the frequency and duration of treatment in a study were frequently missing together. Finally, we identified some variables whose missingness appeared to be related to the size and standard errors of effect size estimates (e.g., Group 1 hours of treatment per week and treatment duration in days), which suggests that missingness was not MCAR.

How to proceed from an EMA will depend on what is known about the data collection and missingness. Based on our analysis, we would be cautious of using analysis methods that assume data are MCAR, such as complete case analysis or shifting units of analysis. Tanner-Smith *et al.* used the EM algorithm to estimate their meta-regression models, which assumes data are MAR. This is consistent with our findings, and it is a common assumption made in analyses of incomplete data. However, we note that distinguishing between data that are MAR or MNAR will often be impossible. We cannot conclude data are MAR or MNAR based on the EMA alone, and it seems infeasible to verify MNAR or MAR without some knowledge of the unobserved data.

The methodology discussed in this tutorial could be used to create different visualizations than were presented in this paper. Our complementary material develops on these results with a vignette that contains further visualizations and executable code implemented in the R computing language. Moreover, even though the data on substance abuse interventions for adolescents has a particular structure with information at the effect size level for each study, the tools exposed in this tutorial can be easily applied to other dataset structures.

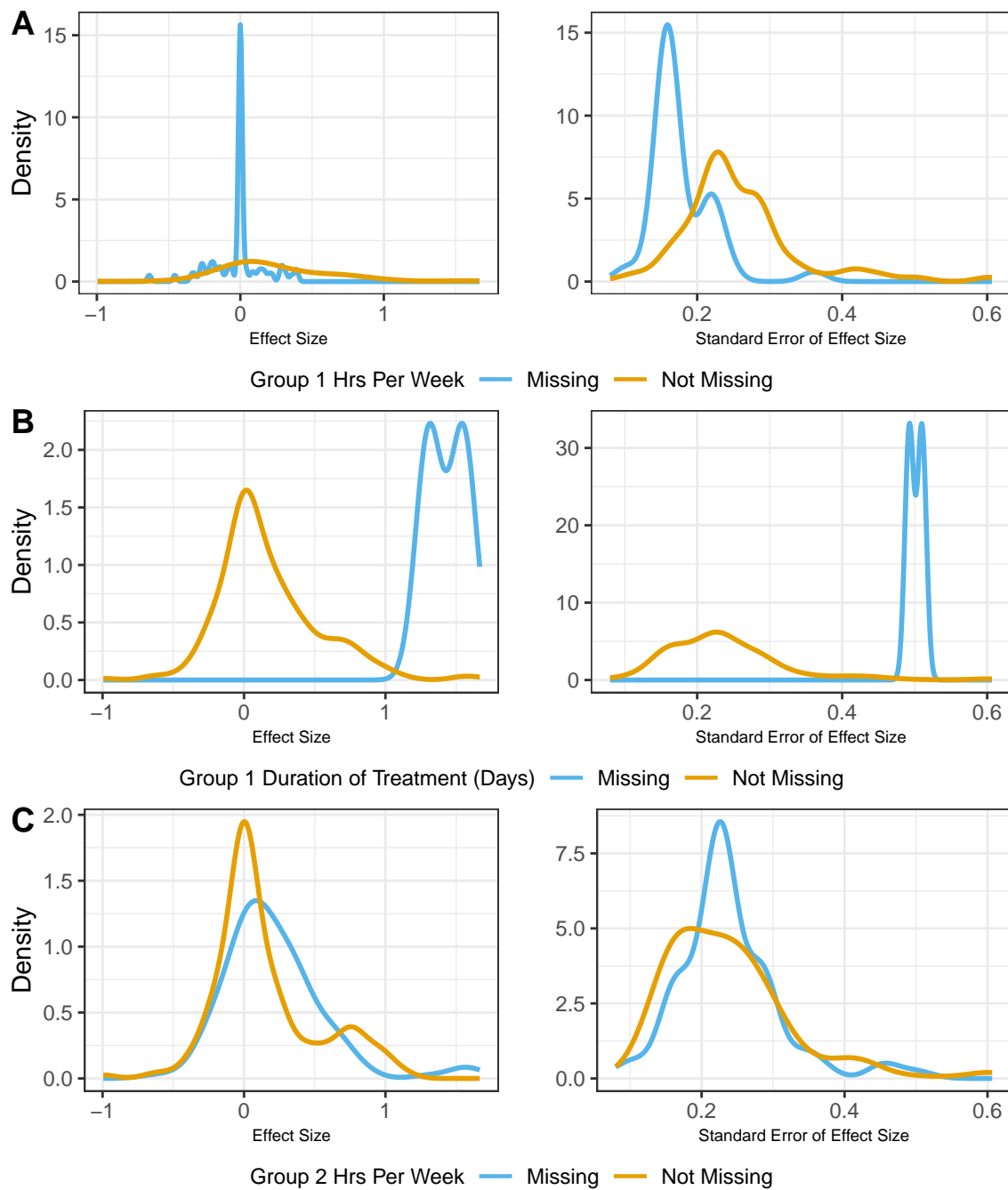


Figure 5: Plot (A) shows that the covariance duration of treatment (days) for group 1 is mostly missing for larger effect size values. Further, the effect size has larger standard error, when this covariate is missing. Plot (B) illustrates a case where the effect size tends to be closer to zero when a particular covariate is missing. Specifically, when treatment contact (hours per week) for group 1 is missing, both the effect size and its standard errors tend to be smaller than when the covariate is present. Plot (C) shows that both, the effect size and its standard errors, have a similar distribution either when the covariate treatment contact (hours per week) for group 2 is present or not.

References

- Baker SG, Rosenberger WF, Dersimonian R. (1992) Closed-form estimates for missing counts in two-way contingency tables. *Statist Med* **11**: 643–657.
- Borenstein M. (2009) *Introduction to meta-analysis*. Chichester, U.K.: John Wiley & Sons.
- Borenstein M, Hedges LV, Higgins JP, Rothstein H. (2012) Comprehensive Meta-Analysis Version 3.0. [https://www.meta-analysis.com/downloads/Meta-Analysis Manual V3.pdf](https://www.meta-analysis.com/downloads/Meta-Analysis%20Manual%20V3.pdf).
- Buja A, Cook D, Swayne DF. (1996) Interactive high-dimensional data visualization. *Journal of Computational and Graphical Statistics* **5**: 78.
- van Buuren S. (2018) *Flexible Imputation of Missing Data, Second Edition*, 2nd edn. Second edition. | Boca Raton, Florida : CRC Press, [2019] | : Chapman and Hall/CRC.
- Cheng X, Cook D, Hofmann H. (2015) Visually exploring missing values in multivariable data using a graphical user interface. *J Stat Soft* **68**.
- Conway JR, Lex A, Gehlenborg N. (2017) UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics* **33**: 2938–2940.
- Cooper HM, Hedges LV, Valentine JC (eds). (2019) *Handbook of research synthesis and meta-analysis* 3rd edition. New York: Russell Sage Foundation.
- Dempster AP, Laird NM, Rubin DB. (1977) Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)* **39**: 1–22.
- Diggle P, Kenward MG. (1994) Informative drop-out in longitudinal data analysis. *Applied Statistics* **43**: 49.
- Egger M, Smith GD, Schneider M, Minder C. (1997) Bias in meta-analysis detected by a simple, graphical test. *BMJ* **315**: 629–634.
- Graham JW. (2009) Missing data analysis: Making it work in the real world. *Annu Rev Psychol* **60**: 549–576.
- Graham JW. (2012) *Missing Data*. New York, NY: Springer New York.
- Graham JW, Cumsille PE, Elek-Fisk E. (2003) Methods for handling missing data. In I.B. Weiner (ed), *Handbook of Psychology*, p. wei0204. Hoboken, NJ, USA: John Wiley & Sons, Inc.
- Hedges LV. (1984) Estimation of effect size under nonrandom sampling: The effects of censoring studies yielding statistically insignificant mean differences. *Journal of Educational Statistics* **9**: 61.
- Hedges LV, Olkin I. (1985) *Statistical methods for meta-analysis*. Orlando: Academic Press.
- Higgins JP, White IR, Wood AM. (2008) Imputation methods for missing outcome data in meta-analysis of clinical trials. *Clinical Trials* **5**: 225–239.
- Ibrahim JG. (1990) Incomplete data in generalized linear models. *Journal of the American Statistical Association* **85**: 765–769.
- Ibrahim JG, Lipsitz SR, Chen M-H. (1999) Missing covariates in generalized linear models when the missing data mechanism is non-ignorable. *J Royal Statistical Soc B* **61**: 173–190.
- Light RJ, Pillemer DB. (1984) *Summing up: the science of reviewing research*. Cambridge, Mass: Harvard University Press.
- Little RJA. (1988) A test of missing completely at random for multivariate data with missing values.

- Journal of the American Statistical Association* **83**: 1198–1202.
- Little RJA, Rubin DB. (2002) *Statistical Analysis with Missing Data*. Hoboken, NJ, USA: John Wiley & Sons, Inc.
- Molenberghs G, Beunckens C, Sotito C, Kenward MG. (2008) Every missingness not at random model has a missingness at random counterpart with equal fit. *J Royal Statistical Soc B* **70**: 371–388.
- Molenberghs G, Kenward MG, Lesaffre E. (1997) The analysis of longitudinal ordinal data with nonrandom drop-out. *Biometrika* **84**: 33–44.
- Newbury-Birch D, Ferguson J, Landale S, et al. (2018) A systematic review of the efficacy of alcohol interventions for incarcerated people. *Alcohol and Alcoholism* **53**: 412–425.
- Pigott TD. (2001a) Missing predictors in models of effect size. *Eval Health Prof* **24**: 277–307.
- Pigott TD. (2001b) A review of methods for missing data. *Educational Research and Evaluation* **7**: 353–383.
- Pigott TD. (2019) Handling missing data. In Harris Cooper, Larry V. Hedges & Jeffrey C. Valentine (eds), *The Handbook for Research Synthesis and Meta-analysis*, 3rd edn. New York: Russell Sage.
- Ramsey AT, Satterfield JM, Gerke DR, Proctor EK. (2019) Technology-based alcohol interventions in primary care: Systematic review. *J Med Internet Res* **21**: e10859.
- Rhoads CH. (2012) Problems with tests of the missingness mechanism in quantitative policy studies. *Statistics, Politics, and Policy* **3**.
- Rosenthal R. (1979) The file drawer problem and tolerance for null results. *Psychological Bulletin* **86**: 638–641.
- Rothstein H, Sutton AJ, Borenstein M (eds). (2005) *Publication bias in meta-analysis: prevention, assessment and adjustments*. Chichester, England ; Hoboken, NJ: Wiley.
- Rubin DB. (1976) Inference and missing data. *Biometrika* **63**: 581–592.
- Schafer JL, Graham JW. (2002) Missing data: our view of the state of the art. *Psychol Methods* **7**: 147–177.
- Seaman S, Galati J, Jackson D, Carlin J. (2013) What is meant by ‘missing at random’? *Statist Sci* **28**: 257–268.
- Tanner-Smith EE, Steinka-Fry KT, Kettrey HH, Lipsey MW. (2016) Adolescent substance use treatment effectiveness: A systematic review and meta-analysis. Office of Justice Programs.
- Tanner-Smith EE, Wilson SJ, Lipsey MW. (2013) The comparative effectiveness of outpatient treatment for adolescent substance abuse: A meta-analysis. *Journal of Substance Abuse Treatment* **44**: 145–158.
- Tierney NJ. (2017) visdat: Visualising whole data frames. *JOSS* **2**: 355.
- Tierney NJ, Cook DH. (2018) Expanding tidy data principles to facilitate missing data exploration, visualization and assessment of imputations. *arXiv:180902264 [stat]*.
- Tipton E, Pustejovsky JE, Ahmadi H. (2019a) A history of meta-regression: Technical, conceptual, and practical developments between 1974 and 2018. *Res Syn Meth* **10**: 161–179.
- Tipton E, Pustejovsky JE, Ahmadi H. (2019b) Current practices in meta-regression in psychology, education, and medicine. *Res Syn Meth* **10**: 180–194.
- Trikalinos T. (2012) OpenMetaAnalyst: Powerful open-source software for meta-analysis. <https://effectivehealthcare.ahrq.gov/products/open-meta-analyst/abstract>.

- Troxel AB, Harrington DP, Lipsitz SR. (2002) Analysis of longitudinal data with non-ignorable non-monotone missing values. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **47**: 425–438.
- Tukey JW. (1962) The future of data analysis. *Ann Math Statist* **33**: 1–67.
- Viechtbauer W. (2010) Conducting meta-analyses in R with the metafor package. *J Stat Soft* **36**.
- White A, Kavanagh D, Stallman H, et al. (2010) Online alcohol Interventions: A systematic review. *J Med Internet Res* **12**: e62.
- Yuvaraj K, Elias SK, Gokul S, Manikandanesan S. (2019) Effectiveness of workplace intervention for reducing alcohol consumption: A systematic review and meta-analysis. *Alcohol and Alcoholism* **54**: 264–271.