HOW TO IDENTIFY AND

MANAGE BIASED STUDIES

by Simon Moss

Introduction

Imagine you are conducting a systematic review to explore whether mindfulness meditation diminishes the frequency of colds. Suppose you have clarified your research objectives, collected the studies that you want to review, and extracted the data from these studies. Before you write your report, you need to evaluate the degree to which the studies are susceptible to biases and then manage these biases somehow.

## Quality instruments

To appraise each study, most researchers utilize a quality instrument—in essence, a sequence of questions the researchers answer to identify biases. Various organizations have developed instruments to evaluate particular study designs. For example, some instruments are primarily designed to evaluate randomized control trials. Other instruments are designed to evaluate quasi-experimental designs, correlational designs, or qualitative designs.

This table presents a set of questions that could be used to evaluate whether studies are biased. The first column presents the questions. The second column illustrates some practices that epitomize high levels of bias. The third column illustrates some practices that epitomize no or low levels of bias. In practice, you would not utilize all these items, but only a subset of questions that seem relevant to your research.

|  |  |  |
| --- | --- | --- |
| **Question or item to assess a source of bias** | **Practices that epitomize**  **a high risk of bias** | **Practices that epitomize**  **a low or negligible risk of bias** |
| **Sample bias** |  |  |
| Is this sample likely to diverge appreciably from the target population? If the target population was not stated explicitly, assume the results are supposed to generalize to the general population |  |  |
| Is this difference between the sample and target population likely to affect the key results | Yes, the sampling bias might compromise the validity of these results | No |
| **Allocation bias** |  |  |
| How, if at all, were participants allocated to conditions? | Participants were not allocated to conditions.  Some criterion, such as birth date or name, determined allocation, but this criterion was not controlled statistically. | A procedure, such as a computer, coin toss, or dice, generated a random number that determined the allocation  Some criterion, such as birth date or name, determined allocation, but this criterion was controlled statistically—called a regression discontinuity design |
| Were the researchers who interacted with the participants or administered the intervention aware of the condition in which these participants were allocated | Yes, researchers could potentially determine the condition in which participants were allocated | No, researchers were blind to the condition in which participants were allocated |
| Were the participants aware of the condition in which they were allocated | Yes | No |
| **Performance bias** |  |  |
| If participants were allocated to conditions, were individuals in each condition treated precisely the same, besides the key intervention the researchers wanted to test | No | Yes  No, but these differences were measured and controlled statistically |
| If a correlation study was utilized, did the researcher statistically control all variables that are likely to affect both the predictors and outcome measures—called spurious variables | No | Yes |
| **Outcome bias** |  |  |
| Were the researchers who administered the measures aware of the condition in which participants were allocated | Yes | No |
| **Attrition bias** |  |  |
| If participants were allocated to conditions, was the withdrawal rate similar in each condition. | Yes, and the participants who withdrew diverged from the participants who did not withdraw on one or more key measures. Thus, withdrawal is not likely to be random. | Yes  No, but the participants who withdrew did not diverge from the participants who did not withdraw on key measures. Thus, withdrawal is likely to be random |
| **Measurements** |  |  |
| Have all the measures demonstrated convergent and divergent validity in this population. That is, do these measures correlate highly with related, but not too highly with distinct, measures | No | Yes |
| **Analysis** |  |  |
| Did the study include too many post hoc tests, increasing the likelihood of false alarm | Yes | No |
| If the researchers measured participants at baseline before the intervention, did the measures differ between the conditions | Yes, and this difference was not controlled statistically | No  Yes, but difference was controlled statistically |
| **Qualitative research** |  |  |
| Were the research procedures—to collect, analyze, and interpret data— documented as precisely as possible? | No | Yes |
| Was the sample and recruitment of participants described clearly | No | Yes |
| Did the researcher clarify and justify the theoretical perspective or philosophy that underpins the research? | No | Yes |
| Can the researcher readily connect each conclusion to tangible data | No | Yes |
| Are the data rich, detailed, and nuanced? | No | Yes |
| Has the researcher offered insights into the diversity of perspectives across participants or circumstances? | No | Yes |

Many of these questions were derived from Crombie (1996), Fink (2005), Greenhalgh (2000), and Petticrew and Roberts (2005). Thus, in your report, while justifying this procedure, you could cite these sources. You could also cite the Cochrane Reviewers’ Handbook, especially if your review is restricted to randomized control trials.

## How should I use these measures of bias?

Although researchers have developed sophisticated techniques to measure and catalogue biases, controversies over how to manage and respond to these biases persist. No standard and satisfactory approach has surfaced.

One controversy is whether researchers should convert this information about biases into some rating, such as something called the Jadad scale for clinical trials. For example, you could simply count the number of features in each study that represent a high risk of bias. Or you could utilize some more sophisticated algorithm. Unfortunately, during recent years, scholars have refuted the utility of this approach. Previous attempts to rate studies have been shown to be invalid or misleading.

Instead, researchers can pursue some other avenues. First, in the table in which you outline each study, you can also summarize the biases. The bottom row in the following table illustrate this approach.

|  |  |  |
| --- | --- | --- |
| Reference | Smith, A. (2018). Mindfulness and infections. Journal of Infections, 1, 1-7. | Jones, B. (2017). Meditation and immunity. Journal of Health, 15, 98-108. |
| N in intervention group | 87 | 132 |
| Withdrawal rate from intervention group | Before treatment: 10%  Before measure: 5%  After measure: 6% | 15% |
| % female in intervention group | 52% | 59% |
| N in comparison group | 86 | 131 |
| % female in comparison group | 48% | 41% |
| Withdrawal rate from comparison group | Before treatment: 6%  Before measure: 2%  After measure: 0% | 8% |
| Age | Mean: 32  Range: 18-60 | Median: 31  Standard deviation: 4.5 |
| Education | 52% Bachelor degree or higher | 28% Bachelor  6% Honors or Graduate Diploma  8% Masters or Doctorate |
| Magnitude of intervention, such as duration | 1 hour a week for 12 weeks | 3 hours a week for 10 weeks |
| Magnitude of comparison, such as duration | 1 hour a week for 12 weeks | 3 hours a week for 10 weeks |
| Effect sizes | Frequency of colds: d = .5 | Duration of colds: d = .23  Frequency of colds: d = .31  Number of absent days ascribed to colds: d = .19 |
| Possible sources of bias | * Sample was younger than population estimates * Participants chose the condition * Administrators of the measures were aware of the condition in which participants were allocated | * The intervention demanded more time than did the control * The researchers who administered the treatment were aware of the condition in which participants were allocated |

In addition, the researchers can specify the impact of excluding studies that demonstrated various biases. For example, they could summarize the key results that were observed before and after excluding studies that were not randomized control designs. Next, they could outline the key results that were observed before and after excluding studies in which the sample was biased and so forth. This approach, comparable to a technique called a sensitivity analysis, offers insight into the effect of these biases on the results. The results that persist regardless of whether biased studies are included or excluded are deemed to be especially robust.

This information could be integrated into the forest plots, display tables, or narrative syntheses. For example, you could write in your results section

|  |
| --- |
| When all studies are included, mindfulness meditation significantly decreased the frequency of colds in 80% of analyses and generated no significant effect in 20% of analyses. However, when the studies were restricted to randomized control trials, mindfulness meditation significantly decreased the frequency of colds in 100% of analyses. |

## References

Cochrane Collaboration (2003). Cochrane Reviewers’ Handbook. Version 4.2.1.

Crombie, I. K. (1996). The pocket guide to appraisal BMJ Books.

Fink, A. (2005) Conducting research literature reviews. From the internet to paper. Sage Publication.

Greenhalgh, T. (2000) How to read a paper: The basics of evidence-based

medicine. BMJ Books.

Petticrew, M., & Roberts, H. (2005). Systematic reviews in the social sciences: A practical guide. Blackwell Publishing