Predictors of chronic pain and level of physical function in total knee arthroplasty: a protocol for a systematic review and meta-analysis — supplementary statistical methods

Introduction

This document is a supplement to the methods described by the protocol by Olsen et al. 2020. It provides more detail on the quantification of associations between predictors and outcomes for use in meta-analysis. The document describes how various measures of association can be placed on a common scale suitable for meta-analysis. While no explicit consideration is given to precision (e.g., confidence intervals and standard errors), standard methods can be used to convert statements of precision between the scales considered.

Methods

Included studies may quantify the association between predictors and outcomes (pain or function) in diverse ways, which makes meta-analysis more challenging than in the typically more uniform case of meta-analyses of the relative safety or efficacy of interventions. In the most general case, for a given predictor, there may be variation between studies with respect to:

- The level of measurement on which the predictor is measured (i.e., predictors may be continuous or categorical). Categorical predictors are usually but not always dichotomous.
- The level of measurement on which the outcome is measured. As for predictors, outcomes may be continuous or categorical.
- The way in which categories are defined. For example, studies may apply different thresholds to underlying continuous variables.
- The way in which associations between predictors and outcomes are measured. For example, studies may report associations as odds ratios, risk ratios, regression coefficients, or correlation coefficients.
- The direction of association, which may vary between studies with respect to both predictors and outcomes. For example, one study may report an odds ratio where being female is the base case while another study may use being male as the base case. Similarly, one study may model pain while another may model lack of pain. This form of variation can be trivially addressed by defining a canonical direction and adjusting the reported estimates such as to enforce the canonical direction of association.

We will be to follow typical meta-analytical practice by analyzing estimates of association on a common scale with a canonical direction of association. Specifically, we will measure association using the correlation coefficient, ρ, because it is defined to be invariant under linear transformations of predictor and outcome variables and under reasonable assumptions can be computed for all combinations of dichotomous and continuous predictors and outcomes (see the following sections). This permits estimates of association to be pooled, even if the included studies measured predictors and outcomes on different levels or scales of measurement, and if

those scales used different units of measurement. Correlation coefficients can be meta-analyzed via transformation to Fisher's *z* (Borenstein 2009). There may be important differences between studies in terms of how predictors and outcomes were measured, which we will model as a source of heterogeneity using random effects meta-analysis (see main protocol text).

Including estimates of association is more challenging still when predictors or outcomes are categorical and there are more than two categories (e.g., if body mass index is categorized as normal, overweight, and obese). Such analyses will typically define a base category and present an estimate of association for each of the other categories compared to the base case (e.g., a risk ratio for postsurgical pain for obese patients versus those of normal weight). Where the categories correspond to an ordinal variable, we will extract and meta-analyze associations for the most extreme case (e.g., for obese versus normal weight patients, rather than overweight versus normal weight); otherwise we will choose a consistent category across studies.

The following table considers four possible combinations of levels of measurement for predictors and outcomes, and four measures of association that we anticipate included studies will report. Numbers in the table correspond to the points that follow the table. Note that while risk ratio (RR) and odds ratio (OR) appear in the same column, we do not assume they are the same quantity (see appendix 1 for how we will treat the distinction). The appendices present equations that relate the various measures of association.

Predictor	Outcome	RR or OR	β	ρ	Notes on units of measurement
Dichotomous	Dichotomous	1	2		
Dichotomous	Continuous		3		Outcome units may vary across studies.
Continuous	Dichotomous	4	5		Predictor units may vary across studies.
Continuous	Continuous		6	7	Units may vary across studies.

RR = Risk ratio; OR = odds ratio; β = linear model coefficient; ρ = correlation coefficient.

1. If the predictor and outcome are both dichotomous and a RR is reported, the corresponding OR can be imputed via equation 1. If an OR is reported or can be imputed, the corresponding correlation coefficient can be obtained via equation 2.
2. If the predictor and outcome are both dichotomous, β is likely to correspond to a log odds ratio or a log risk ratio. This case is essentially identical to case 1.
3. If the predictor is dichotomous and the outcome is continuous, studies are likely to report a difference in means between levels of the predictor. Here, β is a linear model coefficient. The corresponding correlation coefficient can be obtained via equation 3.
4. If the predictor is continuous and the outcome is dichotomous, a RR or OR is likely to correspond to the relative increase in risk or odds of the outcome associated with a one unit increase in the predictor. If a RR is reported, the corresponding OR can be imputed via equation 1. Given an OR, the correlation coefficient (between $\beta = \log \mathrm{OR}$ and the outcome on the logit odds scale) can be computed via equation 3.
5. If the predictor is continuous and the outcome is dichotomous, β is likely to correspond to a log odds ratio or a log risk ratio. This case is essentially identical to case 4.

6. If the predictor and outcome are both measured on continuous scales, β is a linear regression coefficient. The corresponding correlation coefficient can be computed via equation 3.
7. If the predictor and outcome are both measured on continuous scales, ρ is a correlation coefficient and can be used directly.

Appendix 1 — Odds ratios and risk ratios

Odds ratio (OR) is related to risk ratio (RR) via the following equation (Zhang and Yu 1998):

$$\mathrm{RR} = \frac{\mathrm{OR}}{(1 - P_0) + (P_0\ \mathrm{OR})} \quad (1)$$

where P_0 is the baseline risk under the reference condition. Because odds and risks have different definitions, odds ratio is not the same as risk ratio. However, the equation shows that as the baseline risk tends towards zero, RR is increasingly well approximated by OR. Zhang and Yu suggested this approximation is acceptable if baseline risk is below 10%. Because baseline risk of pain after total knee arthroplasty is about 20%, we will not assume that RR and OR can be used interchangeably. When imputing OR from RR, we will assume a baseline risk of 20%.

Appendix 2 — Odds ratios and correlation coefficients

When the predictor and outcome are both dichotomous, the (tetrachoric) correlation coefficient can be computed from an odds ratio via the following equation (Pearson 1900):

$$\rho = \cos\frac{\pi}{1 + \sqrt{\mathrm{OR}}} \quad (2)$$

This coefficient corresponds to the correlation between two continuous variables dichotomized around their means. While the median is often used to dichotomize predictors (and in principle any threshold could be chosen), the mean is a good approximation to the median provided the underlying distribution is not highly skewed. Variation in the definition of thresholds will contribute heterogeneity, which we will model via random effects. This and related methods are discussed in Bonett 2007.

Appendix 3 — Linear model coefficients and correlation coefficients

When the outcome is continuous and the predictor is either dichotomous or continuous, a linear model coefficient β represents the difference in the outcome associated with a change in the predictor (either from one category to the other, or of one unit on a continuous scale). In the general case in which β is adjusted for other predictors, the (partial) correlation coefficient can be computed via the following equation (Freund 2010):

$$\rho = \frac{t}{\sqrt{t^2 + (n - m - 1)}} \quad (3)$$

where $t = {}^{\beta}/_{\mathrm{SE}(\beta)}$, $\mathrm{SE}(\beta)$ is the standard error on β, and *n* and *m* are the sample size and number of predictors in the model, respectively. When an odds ratio is reported (or can be imputed) for a continuous predictor, the linear model coefficient $\beta = \log \mathrm{OR}$ can be obtained and hence a correlation coefficient (between the predictor and the outcome on the logit scale).

Appendix 4 — Fisher's z

Assuming normality, correlation coefficients can be meta-analyzed via transformation to Fisher's *z* (Fisher 1915, Fisher 1921, Borenstein 2009) as computed via:

$$z = \frac{1}{2}\log_e\left(\frac{1+\rho}{1-\rho}\right) = \operatorname{arctanh}\rho \quad (4)$$

References

Bonett, D. G. (2007). Transforming odds ratios into correlations for meta-analytic research. American Psychologist, 62(3), 254–255.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). Introduction to Meta-Analysis. John Wiley & Sons, Ltd.

Fisher, R. A. (1915). Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. Biometrika, 10(4), 507-521.

Fisher, R. A. (1921). On the "probable error" of a coefficient of correlation deduced from a small sample. Metron, 1, 1-32.

Freund, J.F., Wilson, W. J., and Mohr, D. L. (2010). Statistical Methods (3rd Ed.). Academic Press.

Pearson, K. (1900). Mathematical Contributions to the Theory of Evolution. VII. On the Correlation of Characters not Quantitatively Measurable. Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character, 195, 1-405.

Zhang, J. & Yu, K. F. (1998). What's the relative risk?: A method of correcting the odds ratio in cohort studies of common outcomes. Jama, 280(19), 1690-1691.