**Advancing Research Synthesis: A Multilevel and Multivariate Approach to Dependent Effect Sizes**

## Introduction

**Mahmood Ahmad, Royal Free Hospital**

Traditional pairwise meta-analysis relies on the fundamental assumption of independence between effect sizes. However, in contemporary social, behavioral, and medical research, this assumption is frequently violated. Studies often report multiple outcomes, diverse time-points, or multiple treatment arms compared to a single control group, creating a nested data structure where effect sizes are clustered within studies. Ignoring this dependency by averaging effects or selecting a single outcome per study results in a loss of information and statistical power. Conversely, treating dependent effects as independent leads to underestimated standard errors and inflated Type I error rates (false positives). To address these challenges, the 786-MIII Synthesis engine implements a three-level random-effects model, providing a rigorous framework for modern research synthesis.

## Methodology

The core of the application is built upon a three-level hierarchical model estimated via Iterative Generalized Least Squares (IGLS). This approach explicitly partitions heterogeneity into three distinct components: Level 1 accounts for sampling variance (assumed known); Level 2 captures within-study heterogeneity (variation between different outcomes in the same study); and Level 3 captures between-study heterogeneity. By modeling the covariance matrix with a block-diagonal structure, the software accurately reflects the intraclass correlation inherent in clustered data, allowing researchers to synthesize all available evidence without violating statistical assumptions.

To ensure robustness, particularly in datasets with a limited number of studies ($k < 40$), the tool integrates Robust Variance Estimation (RVE) with small-sample corrections. RVE provides valid standard errors and confidence intervals even if the correlation structure between dependent effects is not perfectly specified. Furthermore, the implementation of Cluster Wild Bootstrapping offers a "gold standard" for inference, resampling residuals at the cluster level to generate empirical p-values that are resilient to small-sample bias.

## Conclusion

By combining multilevel modeling with advanced sensitivity techniques like RVE and bootstrapping, the 786-MIII Synthesis tool bridges the gap between complex statistical theory and applied research. It allows methodologists to quantify heterogeneity at different levels (via $I^2$ decomposition) and estimate prediction intervals for future studies, ensuring that meta-analytic conclusions are both statistically valid and clinically meaningful.