The authors respond:

(Response to 15-0611 Schmidt)

Charles Poole,
Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, North Carolina

Ian Shrier,
Centre for Clinical Epidemiology, Lady Davis Institute for Medical Research, Jewish General Hospital, McGill University, Montreal, Quebec, Canada

Peng Ding, and
Departments of Epidemiology and Statistics, Harvard University, Cambridge, Massachusetts

Tyler VanderWeele
Departments of Epidemiology and Biostatistics, Harvard T.H. Chan School of Public Health, Harvard University, Boston, Massachusetts

Theoretical and Empirical Faces of Heterogeneity

We thank Schmidt et al. for contributing their recent paper (1) and letter (2) to the literature on heterogeneity of difference and ratio effect measures. The letter wisely emphasizes the relevance of structural limits on the risk difference (RD). Indeed, if RD = −0.30 among women and unexposed men have a risk of 0.20, RD homogeneity is impossible.

The simulation results (1) for the Q test statistic would be most relevant to the meta-analytic evidence we reviewed (3). With small sample sizes or small fractions of doubly exposed individuals, test size was greater than alpha in the RD homogeneity tests and less than alpha in the tests of odds ratio (OR) homogeneity. This disparity would push the power to detect heterogeneity upward for the RD tests and downward for the OR tests, a result consistent with our conjecture that differential power might be part of the explanation.

Two of us (PD, TJV) are engaged in related methodologic work that has mathematically assessed the three-dimensional volume in the four-dimensional space (of the four outcome probabilities under the different exposure combinations) for which homogeneity holds on the risk difference, risk ratio, and odds ratio scales. The respective volumes are 1.33 for the risk difference, 1.76 for the risk ratio, and 2.47 for the odds ratio scale (4). Thus, relatively speaking, there are more values of the outcome probabilities for which odds ratio homogeneity holds than risk ratio homogeneity and more values for which risk ratio homogeneity holds than risk difference homogeneity. However, these are simply mathematical statements over all possible outcome probabilities. The real question is what

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

values the outcome probabilities take empirically with actual exposures and outcomes and how closely these probabilities approximate homogeneity on different scales.

We are thus heartened by the agreement on the part of Schmidt et al. (1) that the ultimate question is empirical and not theoretical. The suggestion to shun the RD has been made in the belief that it is usually much more heterogeneous than ratio measures such as the OR in empirical research settings. Although there are more heterogeneous possibilities for the RD than for the OR, it would be difficult to defend the assumption that each of those possibilities has the same probability, within or across the many studies that are actually conducted. As noted in our article (3), further evidence is therefore required before concluding that the risk difference is in fact a more heterogeneous measure.

References

1. Schmidt AF, Dudbridge F, Groenwold RHH. Comments on: "Is the risk difference really a more heterogeneous measure?" (letter). Epidemiology. 27(3):xxx.

2. Schmidt AF, Groenwold RHH, Knol MJ, Hoes AW, Nielen M, Roes KCB, de Boer A, Klungel OH. Exploring interaction effects in snmall samples increases rates of false-positive and false-negative findings: results from a systematic review and simulation study. J Clin Epidemiol. 2014; 67:821–829. [PubMed: 24768005]

3. Poole C, Shrier I, VanderWeele TJ. Is the risk difference really a more heterogeneous measure? Epidemiology. 2015; 26:714–718. [PubMed: 26196684]

4. Ding, P., VanderWeele, TJ. The differential geometry of homogeneity spaces across effect scales. arXiv: http://arxiv.org/abs/1510.08534