Quality of Evidence: Analysing GRADE

September 1, 2017

Decisions in medicine are to be made on the best available evidence. To arrive at such a judgement, evidence from clinical trials must be assessed for how good it is. To do so, ranking systems and assessment procedures have been produced by multiple organisations that typically provide guidance on what type of methodology generates the best evidence to support clinical decisions. A prominent assessment procedure is the grading of recommendations assessment, development and evaluation (GRADE) system. GRADE offers a nuanced way of assessing evidence that goes beyond ranking just by methodology and looks at the strengths and weaknesses of individual studies, the data they produce, and any outside factors that may influence the result. GRADE denote how good evidence is by ranking in terms of *Quality of Evidence*. In a recent article (Hultcrantz *et al.* 2017) *Quality of Evidence* is stated as referring to the same concept as *Certainty of Evidence*. In that, a rating of quality "reflect[s] the extent of our confidence that estimates of effect are correct", and as such a high-quality rating indicates how confident one should be that effect estimates are in a range that is as close as possible to the true effect. What then gives one warrant for holding a certain credence in the truth of some result?

In this paper, I show how this concept of Quality (or Certainty) of Evidence is analysable into 3 other concepts that each give support to the overall level of certainty in the truth of the evidence. These concepts are: *Reliability of Methodology*; *Quality of Implementation*; *Quality of the Data*. Reliable methodologies output a high ratio of true results and as such we can have high certainty in the truth of any evidence produced by such methodologies. However, assessment on reliability is only performed relative to idealised methodology, whereas in practice studies are carried out imperfectly. *Quality of Implementation* and *Quality of Data* both assess for these imperfections, and any inadequacies will lower the warrant for being certain in the truth of results. By analysing into these distinct concepts, we are able to fashion an explanation for why we can hold a degree of certainty in the truth of *any* piece or body of evidence. Relying on just the one concept of *Quality of Evidence* does not provide the clarity needed to do this.

The importance of the analysis is two-fold; 1) it affords us a greater understanding of the reasoning behind assessments of evidence in clinical research by being clear on what all the concepts of assessment are; 2) the explanation fashioned from it provides a basis

for extending assessment to other kinds of evidence that are not so highly regarded in clinical research. Evidential pluralism has been argued (see Clarke *et al.* 2014) to be a critical way in which medical practice can be improved, and I sketch out how future work on providing standards for mechanistic evidence in particular can build on the central claims of this paper.

References

Hultcrantz, M., Rind, D., Akl, E. A., Treweek, S., Mustafa, R. A., Iorio, A., Alper, B. S., Meerpohl, J. J., Murad, M. H., Ansari, M. T., Katikireddi, S. V., Östlund, P., Tranæus, S., Christensen, R., Gartlehner, G., Brozek, J., Izcovich, A., Schünemann, H. and Guyatt, G. (2017) 'The GRADE Working Group clarifies the construct of certainty of evidence', Journal of Clinical Epidemiology, 46. doi: 10.1016/j.jclinepi.2017.05.006.

Clarke, B., Gillies, D., Illari, P., Russo, F. and Williamson, J. (2014) 'Mechanisms and the Evidence Hierarchy', Topoi, 33(2), pp. 339–360. doi: 10.1007/s11245-013-9220-9.