

Forthcoming (in press) at Behavioral and Brain Sciences

A disanalogy with RCTs and its implications for second-generation causal knowledge

Authors:

Dr. Kate E. Lynch^{1,2}

¹Department of Philosophy, University of Sydney, Australia

²Charles Perkins Centre, University of Sydney, Australia

Correspondence: kate.lynch@sydney.edu.au

Dr. Rachael L. Brown

Centre for Philosophy of the Sciences, School of Philosophy, Australian National University

Jeremy Strasser

School of Philosophy, Australian National University, Canberra

Dr. Shang Long Yeo

Department of Philosophy, National University of Singapore, Singapore

Abstract:

We are less optimistic than Madole and Harden that family-based GWAS will lead to significant second-generation causal knowledge. Despite bearing some similarities, family-based GWAS and RCTs are not identical. Most RCTs assess a relatively homogenous causal stimulus as a treatment, whereas GWAS assess highly heterogeneous causal stimuli. Thus, GWAS results will not translate so easily into second-generation causal knowledge.

A disanalogy with RCTs and its implications for second-generation causal knowledge

We agree that family-based GWAS are an improvement on traditional GWAS in their ability to rule out confounding common causes. We are, however, sceptical that family-based GWAS will guide research aimed at identifying interventions on non-genetic second-generation variables that can be put to practical use in a manner akin to RCTs. Our scepticism stems from an overlooked disanalogy between family-based GWAS and RCTs—*the heterogeneity of the causal stimulus*—and its impact on non-uniformity.

In most RCTs, individuals in the treatment group receive the same, or as similar as possible, treatment or causal stimulus, such as a drug or educational intervention (*causal stimulus homogeneity*). The same is true of Mendelian Randomization trials (which perhaps inspired Madole and Harden’s arguments)—the causal variable(s) being investigated are relatively homogenous exposures. In contrast, family-based GWAS make claims about the average treatment effects of thousands of genetic variations distributed across a population. This aggregation is assigned a single variable, ‘genes’, which can be demonstrated as causal to an outcome to some degree, but nonetheless shows high *causal stimulus heterogeneity*.

Causal stimulus heterogeneity differs from the *heterogeneity of treatment effects*—a feature of both RCTs and GWAS—, which Madole and Harden discuss in the article. The heterogeneity of treatment effects concerns the non-uniform effects of casual stimuli due to interactions with background factors like physiology and environment. This type of non-uniformity is seen in both GWAS and RCT studies, but is particularly challenging for GWAS because of causal stimulus heterogeneity. In GWAS, the complex role of the environment in the expression of the phenotype is amplified because the causal stimulus is varied and heterogeneous between individuals (Lynch 2021). This non-uniformity means that the associations GWAS uncover between phenotypes and large aggregates of gene variants are very difficult to connect to mechanisms and function (Matthews and Turkheimer 2022). A similar challenge occurs in microbiome research: significant within-population physiological and environmental variation make it difficult to track pathways between microbes and outcomes, limiting the scope for causal inference (Lynch et al. 2019).

Causal stimulus heterogeneity increases non-uniformity and hampers tracing of mechanisms. This is due to variation in nature of ‘the same’ treatment upon subjects. A simple hypothetical illustrates this well. Consider three different drug treatments: First, a drug with a single active ingredient (e.g., lithium). Second, a drug with thousands of ingredients of small efficacy. Third, a drug with thousands of ingredients of small efficacy, where each pill has one ingredient or an alternative at random according to a defined chance procedure. In all three cases, an RCT can determine whether the treatment drug has an average effect compared to a control, and thereby generate first-generation causal knowledge. This is possible even in the face of non-uniformity due to the heterogeneity of treatment effects. The prospect for these results to advance second-generation causal knowledge diminishes, however, across the three cases. The high causal stimulus heterogeneity in the third case is likely to produce non-uniform causal pathways from the very first steps, thus making it difficult or impossible to trace mechanisms from particular drug ingredients given only associations between treatments and outcomes.

A high degree of causal stimulus heterogeneity is typical for GWAS, including family-based ones. To analogise with our hypothetical drug case; the first drug is akin to a single gene cause, the second a specific aggregate of genes, and the third an aggregate of many genes,

which at the individual level is summarised by a polygenic risk-score. Polygenic risk-scores are highly heterogeneous causal stimuli with non-uniform effects that make it extremely difficult to trace mechanisms from particular genetic ingredients in the causal stimulus, in all but the simplest cases of gene expression (where GWAS is unnecessary). Even if we could hold environments fixed between individuals (thereby reducing the potential for non-uniformity due to background conditions), in GWAS there is typically too much variation between individuals in how the causal stimulus works at the ‘lower’ biological level for effective ‘bottom-up’ investigations of intermediate causes through biological mechanisms. In our view, this largely precludes these studies from providing the sort of causal knowledge required to identify mechanisms and intermediaries for investigation with second-generation studies.

A work-around might be Harden’s proposal of Phenotypic Annotation, which rests on the statistical investigation of intermediate causes through mediation analysis (Belsky and Harden 2019; see Belsky et al. 2016 for an example study). Mediation analysis tests variables correlated with the stimulus to determine whether (and to what extent) they mediate the causal paths from stimulus to outcome. Such intermediaries could be possible targets for intervention in second-generation studies. In the simplest case, the genetic stimulus would act as an instrumental variable on the potential intermediary (as in Mendelian Randomization, see Davey & Ebrahim 2003) allowing for measurement of the intermediary’s causal effect. However, mediation analysis is tricky at the best of times (see Pearl 2014 for a principled approach). The possibility of confounding common causes between intermediary and outcome is a serious challenge. Common causes (such as other genetic or environmental causes) may account for the relationship between potential intermediary and outcome. In this case, intervening on the potential intermediary will not cause the outcome. A heterogeneous causal-stimulus, such as a polygenic risk-score, effectively carries a potential common cause within itself: the different ways that the causal stimulus may be realized. The use of an average causal stimulus (by definition) precludes control of this common cause. To determine if a potential intermediary is indeed a cause of the outcome, one would need to do a RCT or another kind of study.

In conclusion, we agree that family-based GWAS provides one kind of causal information that has been missing from traditional heritability and GWAS (see Lynch 2017 for the limitations of causal heritability claims). Unfortunately, the general heterogeneous nature of the genetic variation studied means that this information will not translate easily into second-generation causal knowledge.

Funding Statement

KL was supported under Australian Research Council's Discovery Projects funding scheme (project number FL170100160), RB was supported by funding through the ANU VC Futures Scheme.

Conflicts of Interest Statement

KL, RB, JS, SLY declare no conflict of interest.

References

- Belsky, D. W., & Harden, K. P. (2019) Phenotypic annotation: Using polygenic scores to translate discoveries from genome-wide association studies from the top down. *Current Directions in Psychological Science*, 28, 82–90. doi: 10.1177/0963721418807729
- Belsky, D. W., Moffitt, T. E., Corcoran, D. L., Domingue, B., Harrington, H., Hogan, S., ... & Caspi, A. (2016) The genetics of success: How single-nucleotide polymorphisms associated with educational attainment relate to life-course development. *Psychological Science*, 27, 957-972. doi: 10.1177/0956797616643070
- Davey Smith, G., & Ebrahim, S. (2003) ‘Mendelian randomization’: can genetic epidemiology contribute to understanding environmental determinants of disease? *International journal of epidemiology*, 32(1), 1-22.
- Lynch, K.E. (2017) Heritability and causal reasoning. *Biology & Philosophy*, 32(1) 25-49.
- Lynch, K.E., Parke, E, O'Malley M. (2019) How causal are microbiomes? A comparison with the *Helicobacter pylori* explanation of ulcers. *Biology and Philosophy*, 34: 62.
- Lynch, K.E. (2021) The meaning of cause in genetics. *Cold Spring Harbour Perspectives in Medicine* Eds. George Davey Smith, Rebecca Richmond and Jean-Baptiste Pingault. A0405019.
- Madole, J. W., & Harden, K. P. (2021). Building Causal Knowledge in Behavior Genetics. *Behavioral and Brain Sciences*, 1-76.
- Matthews, L. J., & Turkheimer, E. (2022) Three legs of the missing heritability problem. *Studies in History and Philosophy of Science*, 93, 183-191.
- Pearl, J. (2014). Interpretation and identification of causal mediation. *Psychological methods*, 19(4), 459.