The Problem of Mooted Models for Analyses of Microbiome Causality

Authors: Justin Donhauser*, Sara Worley*, Michael Bradie*, and Juan L. Bouzat**

Target article: "How Causal are Microbiomes? A Comparison with the Helicobacter pylori Explanation of Ulcers" by Kate E. Lynch, Emily C. Parke, Maureen A. O'Malley

Corresponding author: Justin Donhauser (jdonhau@bgsu.edu)

Lynch, Parke, and O'Malley highlight the need for better evaluative criteria for causal explanations in microbiome research. They propose new interventionist criteria, show that paradigmatic examples of microbiome explanations are flawed using those criteria, and suggest numerous ways microbiome explanations can be improved. While we endorse their primary criticisms and suggestions for improvements in microbiome research, we make several observations regarding the use of mooted causal models in microbiome research that have significant implications for their overall argument. In sum, we contend that their critique is too modest and that even flawed causal inferences like those they criticize can be valuable for generating better causal models and evaluating explanatory outcome in individual cases.

1. Problems with Mooted Models

As an alternative to using Koch's postulates for identifying microbiome causal factors, Lynch, Park and O'Malley defend Proportionality, Specificity, and Stability interventionist criteria for evaluating the strength of possible causal explanations. They show that widely received explanations of obesity and mental health outcomes do not fulfill these criteria. Moreover, they suggest that such causal inferences could be improved if microbiome researchers: (i) are more specific about how they conceive of target microbiomes, (ii) pay explicit attention to how well their mooted explanatory models meet their criteria for causal explanations, (iii) define outcomes in more nuanced, non-binary, ways, and (iv) focus on specified factors within target microbiome rather than making claims about any microbiome as a whole. We contend that even adopting these suggestions won't produce causal explanations that fulfill the interventionist criteria they endorse. Examining the cases they use to illustrate the virtues of that interventionist schema helps show why they won't.

1.1. H. pylori and Ulcers

Lynch, Parke, and O'Malley present the *H. pylori* case as a "traditional" microbial causal explanations rather than microbiome causation; since it appeals to one species rather than a

* Department of Philosophy, Bowling Green State University, Bowling Green, OH
** Department of Biological Sciences, Bowling Green State University, Bowling Green, OH

microbiome as a whole. They argue that the claim that '*H. Pylori* causes ulcers' fulfills the Proportionality criterion (or at least does better than the competing hypothesis that ulcers are due to activation of the immune system) even though it fails the Stability or Specificity criteria. The reasons it fails the latter are instructive.

The endorsed interventionist schema requires specifying a mooted model in which values of all variables except the one being tested can remain fixed while the tested one is intervened upon. Importantly, the quality of inferences that can be drawn from experimental interventions depends on the quality of that mooted model. If potentially relevant confounds are left out, the causal inference to *H. pylori* is underdetermined. Lynch, Parke, and O'Malley show that the model according to which '*H. pylori* causes ulcers' does not even "achieve a key explanatory standard in microbiology," since it fails Koch's postulate according to which an acceptable microbial cause must be shown to correspond specifically with the target disease and shown to induce the disease when introduced in healthy animals. They note that the postulate was not even fulfilled in the landmark Marshall *et al* (1985) study in which Marshall drank a culture of *H. pylori* to test himself as an animal host—because he was the sole test host and didn't even go on to develop ulcers but just gastritis. Moreover, key evidence supporting the inference to *H. pylori* was derived from an overly simplistic mooted model and accordingly oversimplified experimental designs.

The mooted model was myopically focused on only one type of bacteria, because it was the one that the researchers found and could culture and because four patients with ulcers who had *H. pylori* in their guts responded to treatment with multiple antibiotics that eliminated *H. pylori* (Marshall 2005). Only *H. pylori* was cultured from the stomachs of ulcer patients. However, when additional species are present, the success of antibacterial treatments does not warrant the inference to *H. pylori*—because such interventions only show that *something* present in the microbiome was causally relevant to ulcers not that it was *H. pylori*. So, in this case, interventions established the causal claim because they only targeted the mooted cause. Of course, a better experimental design would work with a model that accounts for potential confounds and tests for outcomes whether or not they are present.

Lynch, Parke, and O'Malley's discussion reveals similar problems in the case of *C. diff*. Fecal Microbiome Transplant (FMT) interventions on entire microbiomes are effective in curing *C. diff*. Following principles that allow the inference to the causal role of Helicobacter thus supports inferring a causal explanation from this intervention. This, they explain, is in some tension with a more focused explanation of the treatment of *C. diff* in terms of specific organisms and pathways. The model that refers to the entire microbiome seems to fail the Proportionality criterion even though interventionist principles license it. However, the apparent tension between these explanations is resolved once we recognize that good interventionist inferences require sufficiently detailed mooted models. The success of the intervention tells us something in the microbiome is plausibly playing a role, but a more focused intervention is needed to determine what. And to know how to design a more focused intervention, we need more detailed models that specify the variables to be kept fixed and those to be intervened upon.

These cases seem to show that problems with Proportionality derive from inadequate mooted models. Failures of the Specificity and Stability criteria on the other hand have more to do with the complexity of causal relations in nature. A causal relationship's "stability" is a function of its robustness across a variety of background conditions. The relationship between *H. pylori* and ulcers is not very robust, since background conditions play a large role in determining whether infected people contract ulcers. Similarly, Specificity is not satisfied if there are multiple, independent, pathways to an effect. The works Lynch, Parke, and O'Malley reference mention two (*H. pylori* and non-steroidal anti-inflammatory usage). More recent works identify many more independent causes. One large study found "risk for ulcer related to stress was similar among subjects who were *H pylori* seropositive, those who were *H pylori* seronegative, and those exposed to neither *H pylori* nor nonsteroidal anti-inflammatory drugs" " (Levenstein *et al* 2015, 498). Multivariable analysis in that study also showed that "stress, socioeconomic status, smoking, *H pylori* infection, and use of nonsteroidal anti-inflammatory drugs were independent predictors of ulcer" (Ibid.).

1.2 Microbiomes and Obesity

Let us now consider Lynch, Park, and O'Malley's recommendations (i-iv) for improving causal inferences in microbiology. It seems these recommendations are not sufficient to overcome the above sorts of problems, at least in the near term. Consider, for instance, their critique of the mooted relationship between the Firmicutes/Bacteroidetes ratio and obesity on the grounds of Stability and the existence of potential confounds. Here, developing an apt mooted model is impossible in practice even though a widely-received causal explanation is well specified; since the mooted relationship is not between a microbiome taken as a unitary entity but between obesity-risk and the ratio of two phyla in the guts of mice and humans. Despite this specificity, the causal relationship lacks stability because it is sensitive to background conditions—of which some are known (e.g. fat in the diet) while others are not (since increases and decreases in Firmicutes are both associated with obesity). The mooted model also overlooks potential confounds; including the noted problems with germ-free mice, and, alternatively with mice treated with antibiotics. Since we do not yet have a theoretical handle on these background conditions and potential confounds, we are not yet in a position to construct an apt mooted model to guide experimental manipulations that could found strong causal claims.

Given the complexity of interaction effects, we may never arrive at particularly stable causal claims about microbiome factors and obesity-risk. We may only ever be able to provide much more restricted claims; e.g., that certain components of a microbiome, given certain background conditions, tend to produce some degree of obesity-risk. Lynch, Park, and O'Malley note that similar comments apply to Specificity. If, as seems plausible, there are multiple pathways to similar outcomes, the relevant sorts of causal claims fail the Specificity criterion. We cannot say that obesity is *always* produced by a certain feature of the microbiome, but only that sometimes, in some circumstances, obesity is produced by that feature.

Studies on obesity and specific bacterium even show that sensitivity to background conditions makes constructing apt mooted model impossible in practice. Consider studies on correlations

between *H. pylori* infections and obesity for instance. Some find substantive correlations (see, e.g., Chen et al. 2018). Others find that *H. pylori* abundance has no relation to BMI in particular demographics (Kawano *et al* 2001; Kyriazanos *et al 2002*; Archimandritis *et al* 2003). And still others find no relationship at all between *H. pylori* and being overweight (see, e.g., Ioannou *et al 2005;* Cho *et al* 2005).

1.3. Microbiomes and Mental Conditions

The above points all apply equally well to the example of FMT experiments and probiotic manipulations used to establish causal connections between features of microbiomes and mental conditions. Lynch, Parke, and O'Malley acknowledge that both the nature of microbiomes (as a group of interacting microbial species) and the nature of mental health states (as complex traits affected by multiple genes, the environment, and their interactions) contribute to a multifactorial range of possible developmental pathways. They argue, convincingly, that claims about the causal role of any microbiome as a whole are thus unlikely to fulfill the Proportionality or Specificity criteria. Depression and anxiety, they suggest, are multifactorial traits which involve complex interactions between microbiomes, hosts, and their environments. It's therefore unlikely that claims involving entire microbiomes could be specific, since a given microbiome-type is unlikely to be in 1-1 correspondence with a particular psychological profile. We believe it is also unlikely that they will be stable, since a given microbiome-type will probably not produce a specific set of psychological traits across a range of background conditions.

It is also unclear how following Lynch, Parke, and O'Malley's suggestions, and adopting more highly specific accounts of microbiomes or focusing on specific behavioral components, will help overcome either of these problems for microbiome research on mental conditions. Since psychological traits are due to complex interactions between a microbiome, a host (even down to epigenetic artifacts), and their environment, there's just no reason to expect that there will be specific or stable relationships between any given microbiome components and specific mental conditions. There are just too many confounds and background influences to develop apt mooted models to establish such relationships.

2. Flawed Causal Inferences about Microbiomes can be Valuable

We agree that researchers should aspire to explanations that fulfill the Specificity, Stability, and Proportionality criteria. Yet, even explanations that fall short can be, and are often, pragmatically and heuristically useful. For instance, learning that FMT transplants can cause obesity in germ-free mice tells us something about where to look for relevant causal mechanisms, even though it does not fulfill the Stability criterion. Finding out that this effect is moderated by the level of dietary fat also gives us information about which causal models might be plausibly hypothesized. Flawed models can be valuable for developing and refining hypotheses that can help generate more and more refined causal models. We contend that it is

therefore crucially valuable, especially at early stages of a research program, to value crude and flawed hypotheses and models of potential causal mechanisms.

Flawed models can also have value for helping evaluate particular cases and potential courses of disease treatment. We know that in *some* cases, antibacterials and means of changing the composition of gut microbiomes can be used to effectively treat and/or prevent ulcers. In *some* cases, such interventions can help treat and/or prevent obesity. And in *some* cases, they can be used to effectively treat and/or prevent certain mental conditions. Knowing these things is obviously useful in particular cases, even though the mooted models these conclusions derive from are wrong and even if generalizable mooted models are impossible.

In view of the complexity of microbiomes and the broader networks of which they are parts, we believe that it is at this stage practically impossible to produce apt mooted models to found generalizable microbiome explanations. The appeal to more rigorous interventionist explanatory criteria at this stage of our understanding is thus a step in the right direction but at once quixotic. But all is not lost. Interventionist explanatory criteria are no doubt useful and explanatory models that fail interventionist criteria can be crucially valuable nevertheless.

Works Cited

Levenstein, S., Rosenstock, S., Jacobsen, R. K., & Jorgensen, T. (2015). "Psychological Stress Increases Risk for Peptic Ulcer, Regardless of Helicobacter pylori Infection or Use of Nonsteroidal Anti-inflammatory Drugs." *Clinical Gastroenterology and Hepatology, 13*(3), 498-506.e491.

Archimandritis AJ, et al. (2003) The incidence of *Helicobacter pylori* is not increased in obese people, but the risk of duodenal ulcer is higher in *H*. *pylori*-positive people of high obesity. J. Clin. Gastroenterol;36:184–185. doi: 10.1097/00004836-200302000-00022.

Chen, L. W., Kuo, S. F., Chen, C. H., Chien, C. H., Lin, C. L., & Chien, R. N. (2018). A community-based study on the association between Helicobacter pylori Infection and obesity. *Scientific reports*, *8*(1), 10746. doi:10.1038/s41598-018-28792-1

Cho I, et al (2005) *Helicobacter pylori* and overweight status in the United States: data from the Third National Health and Nutrition Examination Survey. Am. J. Epidemiol.;162:579–584. doi: 10.1093/aje/kwi237.

Ioannou GN, Weiss NS, Kearney DJ. (2005) Is *Helicobacter pylori* seropositivity related to body mass index in the United States? Aliment. Pharmacol. Ther.; 21:765–772. doi: 10.1111/j.1365-2036.2005.02369.x.

Kawano S, et al (2001) *Helicobacter pylori* infection does not affect serum leptin concentration and body mass index (BMI) in asymptomatic subjects. J. Gastroenterol.;36:579–580. doi: 10.1007/s005350170064.

Kyriazanos ID, et al. The incidence of *Helicobacter pylori* infection is not increased among obese young individuals in Greece. J. Clin. Gastroenterol. 2002;34:541–546. doi: 10.1097/00004836-200205000-00012.