**Can Intersectionality Theory Enrich Population Health Research?**

Mark A Green1\*†, Clare R Evans2\*, SV Subramanian3,4.

1 Department of Geography & Planning, University of Liverpool, Liverpool, UK.

2 Department of Sociology, University of Oregon, Eugene, OR, USA.

3 Center for Population and Development Studies, Harvard University, Cambridge, MA, USA.

4 Department of Social and Behavioral Sciences, Harvard Chan School of Public Health, Harvard University, Boston, MA, USA.

\* MAG and CRE contributed equally to this work and choose to share first authorship.

† Corresponding author. Address: Department of Geography & Planning, Roxby Building, University of Liverpool, Liverpool, L69 7ZT, UK. Email: mark.green@liverpool.ac.uk.

Keywords: Intersectionality; multi-level modelling; health inequalities; epidemiology.

Originating in black feminist scholarship (Collins 1990, Crenshaw 1989), intersectionality theory is emerging as a cornerstone of sociological thought. It encourages us to consider the ways in which upstream social determinants such as racism, sexism and classism form interlocking systems of oppression that shape the experiences and life chances of individuals as a consequence of their multi-dimensional social identities. Contextual forces such as sexism or racism do not operate in isolation but interact with each other in the production of health inequalities. Intersectionality is being increasingly adopted in social epidemiology because it dovetails with the domain’s focus on the underlying power structures that produce inequalities (rather than inequalities simply resulting from the accumulation of independent risk factors). The most common way in which intersectionality is operationalised by health inequalities researchers has been “inter-categorical intersectionality” (McCall 2005), which calls for considering numerous interactions between dimensions of social identity (and the social forces they are proxies for). In this commentary, we examine the potential contributions of intersectionality to applications in epidemiology and health-related fields. We identify challenges and opportunities for research, and outline future directions.

**Sick Individuals, Sick Populations and Intersectionality**

In their study published in *Social Science & Medicine*, Wemrell, Mulinari and Merlo (2017) embrace intersectionality in order to inform a longstanding debate between so-called “risk factor epidemiology” and “eco-epidemiology” (Susser and Susser 1996a; 1996b). The risk factor approach to epidemiology is concerned with studying the determinants of disease in individual cases. Or put more eloquently by Roy Acheson: “Why did *this* patient get *this* disease at *this* time?” (cited in Rose, 1985, p32). Eco-epidemiology broadens our assessment of the determinants of health through incorporating the wider ‘ecology’ individuals operate their lives within (i.e. we are not just interested in biological factors, but also social context).

Epidemiology has long recognized that the causes of individual *cases* do not necessarily match the causes of *population* level patterns (or differences in average values between populations) (Rose, 1985). For example, as Davey Smith (2011) noted: “most cases of lung cancer are attributable to smoking, but many smokers do not develop lung cancer” (p549). At the population level, smoking accounts for upwards of 90% of the variation in lung cancer rates between populations and within a population over time, however it accounts for less than 10% of the variation between individuals. Indeed, only half of regular smokers will die of smoking-related diseases (Doll et al. 1994). Eco-epidemiology builds from this distinction to explore causes exerting influence simultaneously through multiple pathways at multiple levels — including *within* groups and *between* groups (Susser and Susser, 1996b).

Intersectionality should not be strictly treated within a risk factor framework, both on theoretical grounds and because inter-categorical intersectionality does not fix what is an inherent flaw in the risk factor approach. Including more interacting dimensions of social identity ultimately does not serve to improve our ability to distinguish between individuals who will get sick and individuals who will not. Indeed, Wemrell and colleagues’ results suggest that the introduction of interaction effects adds very little to their ability to predict who will develop ischaemic heart disease. While inter-categorical intersectionality is a useful tool for studying differences between populations, regression models with numerous interaction terms inherently overlook the heterogeneity of risk that remains *within* these identities.

Wemrell and colleagues’ application of discriminatory accuracy within an intersectionality framework provides a novel methodological addition to the anti-categorical intersectionality literature and utilizes an alternative branch of intersectional thought that is fairly novel in health-related research. Anti-categorical intersectionality critiques approaches that categorise individuals on the grounds that social experiences are irreducibly complex and that categorisation usually leads to reification. Discriminatory accuracy allows us to quantify how effective (or ineffective) the categories used in the analysis are at distinguishing between cases and non-cases. Discriminatory accuracy challenges how we categorise risk by providing a formal measure of outcome heterogeneity *within* categories (or identities) and reminding us that social identity categories are not static or inflexible. Given that the majority of anti-categorical work has been qualitative or purely theoretical, discriminatory accuracy is an important tool for quantitative researchers (both intersectional and beyond) wanting to critically evaluate how useful their models are in predicting the incidence of health outcomes.

**Important Cautions for Using Intersectionality**

Intersectionality should not be operationalised merely as the justification for including interaction effects within a regression model. Nor does it naturally align itself with a risk-factor approach, wherein ever narrower and more specific “risky” identities are identified. Intersectionality will be most effective where combined with social theory on the production of health inequalities. To do otherwise is to lose the focus intersectionality theory brings to underlying power structures and social determinants, and to treat the identities themselves as being inherently “risky” rather than as proxies for social position and experience. We argue that greater attention should be paid in research to interactions in general, and that this would not necessarily require allusions to intersectionality theory. However, if researchers choose to employ intersectionality theory to motivate their study design, then they should also commit to follow through when interpreting these interacting dimensions of social identity and the social determinants they represent.

The most obvious question raised by Wemrell and colleagues’ paper is whether we *should* be focusing on improving our ability to perfectly discriminate between cases and non-cases. Achieving 100% discriminatory accuracy may not be feasible (due to cost, availability and the sheer volume of data this may require) or even desirable. While a case could easily be made that understanding the origins of within-group heterogeneity remains one of the great unanswered challenges in epidemiology, we should not neglect population-level analyses which have historically been important for understanding the determinants of inequalities (e.g. Doll et al. 1994). Intersectional approaches are key here since they integrate our understanding of between and within group differences, allowing for consideration of both dimensions for understanding population health.

The results of Wemrell and colleagues’ study may lead some to argue that we should utilize anti-categorical, as opposed to inter-categorical, intersectional approaches. The anti-categorical approach is fundamentally critical of any categorisations — even provisionally and in full recognition of their inherent fluidity and social construction. Yet this is counter to the very nature of what inequalities research (and most of epidemiology) requires — the provisional division of people into groups so that we can measure whether (and why) one group is worse off than another. The fluidity of categories (both between and within) highlights the tension between focusing on individuals (a natural unit of observation) and the ‘population’ (which can be arbitrary). While treating individuals as unique may be conceptually pleasing it is not practical for delivering population-level interventions which may seek to address inequalities. Keeping in mind the importance of not treating intersectional identities as risk factors, but rather as proxies for social forces, the inter-categorical approach fits better with the methods and mission of social epidemiology.

**Toward an Eco-Epidemiologic Approach to Intersectionality**

Current quantitative applications of intersectionality theory have failed to find a truly eco-epidemiologic approach. Such an approach would allow for the simultaneous consideration of *between group* variation (and the causes of these population level inequalities) and *within group* variation (and the causes of within group heterogeneity). Most common approaches utilize a fixed effect regression model with a full complement of interaction parameters between each explanatory variable. However, this conventional approach allows for consideration only of between group differences, and fails to explicitly model the high within group variance that remains unexplained. A multilevel modelling approach to intersectionality research, where individuals (level 1) are nested within their intersectional social identities (level 2), has been recently proposed as one solution (Evans 2015). Such an approach has methodological advantages, as well as being theoretically concordant with eco-epidemiology.

To illustrate the differences between these approaches, consider the simple case of health outcome $y$, which is a continuous measure for individual $i$ who belongs to intersectional social identity category $j$. An intersectional model might consider the interaction between gender (male and female) and race (non-Hispanic black and non-Hispanic white). In practice it is common to consider many more dimensions of identity simultaneously, but this simple case is sufficient for illustration. The conventional fixed effects model would be specified as:

$$y\_{i}=β\_{0}x\_{0i}+β\_{1}x\_{1i}+β\_{2}x\_{2i}+β\_{3}x\_{3i}+e\_{0i}$$

where the fixed effect parameters are dummy variables which taken together fully specify individual $i$’s intersectional identity. $x\_{0i}$ represents the reference group (white males), $x\_{1i}$ represents female (1=yes, 0=no), $x\_{2i}$ represents black (1=yes, 0=no), and $x\_{3i}$ represents being *both* black *and* female (1=yes, 0=no). $e\_{0i}$ is the difference between the group average and the observed outcome for individual $i$. In this modelling approach, finding the interaction term $β\_{3}$ to be statistically significant would be interpreted as an intersectional effect observable for black women.

The multilevel (random effects) model, on the other hand, would be specified as:

$$y\_{ij}=β\_{0}x\_{0j}+β\_{1}x\_{1j}+β\_{2}x\_{2j}+u\_{0j}+e\_{0ij}$$

$$\left[u\_{0j}\right]\~N\left(0,σ\_{u0}^{2}\right)$$

$$\left[e\_{0ij}\right]\~N\left(0,σ\_{e0}^{2}\right)$$

Where the main effects for gender and race are still included in the fixed part, but the interaction parameter is excluded. Replacing it is a hierarchical data structure nesting individual $i$ in social identity $j$. $e\_{0ij}$, which is normally distributed with mean of 0 and variance of $σ\_{e0}^{2}$, represents the difference between the observed outcome for individual $i$ in social identity $j$ and the average outcome in his or her category. $u\_{0j}$, which is normally distributed with mean of 0 and variance of $σ\_{u0}^{2}$, represents the difference between what is predicted for social identity category $j$ and *what might have been predicted for it based purely on the additive contributions of the main effects*. In other words, finding $u\_{0j}$ to be 0 for a particular intersectional category would indicate that that identity has no discernable intersectional effect above and beyond the additive dimensions of the identity. Critically this enables us to examine intersectional effects for all identities, without requiring *post hoc* calculations or specifying *a priori* which identity should be the reference level. Additionally, it enables us to determine the “added value” of considering interactions between dimensions of identity beyond the contributions of the main effects.

When researchers wish to model high-dimension interactions, as in the case of intersectionality, the conventional approach to modelling interactions has limitations when it comes to scalability, model parsimony, reduced sample size in some intersectional identity groups, and interpretability. As the number of dimensions of social identity considered increases, the number of parameters required by the fixed effects model increases *geometrically*. Due to the multi-level approach’s precision weighted estimation along with borrowing strength due to the inclusion of random effects, the approach adjusts for sample size issues which may provide unrealistic results for some groups under the fixed effects approach. Considering additional dimensions of identity in a multilevel approach, however, would increase the number of parameters *linearly* (to account for new main effects, but not new interaction parameters) and it would increase the number of level 2 units (i.e., social identities). This increases the scalability and model parsimony of the multilevel approach over conventional models. Additionally, the reduced number of parameters enables improved interpretability of findings.

The concordance between eco-epidemiology and the multilevel approach to studying interactions and inter-categorical intersectionality is clear. Multilevel models parse the total variance across each ecological level considered—in this case, the between-category ($σ\_{u0}^{2}$) and within-category ($σ\_{e0}^{2}$) levels. This enables a clear presentation of between-group differences in risk that also explicitly acknowledges the low discriminatory accuracy (i.e., high within group heterogeneity of risk) that remains. Exploring possible causes using a multilevel framework would enable clear thinking about which level those causes are operating at—do they explain differences between groups or variation within groups?

**Other Future Directions**

While discriminatory accuracy offers a measure of within-group heterogeneity, it is merely a descriptive tool and does not offer explanations for such heterogeneity. There is growing interest in how measures of health outcomes, social characteristics and health-related behaviours contain a diversity of experiences which are not always recognised (e.g. Green et al. 2016). Understanding the nature of such heterogeneity among cases both within and between groups would enrich both risk-factor and eco-epidemiological approaches to research.

While quantitative applications of intersectionality in epidemiology to date have acknowledged in general terms that variables such as race and sex are proxies for racism and sexism, an important next step is to explore the ways in which these social determinants of health inequalities operate at multiple levels simultaneously. Take for instance the impacts of racism. Racism can be experienced in an internalized way, interpersonally or structurally. In the case of segregation, an individual may experience discrimination not just because they live in a segregated neighbourhood, but because of issues at the wider city, region or national level (Manley et al. 2015). Further consideration should be given to the ways in which the social significance of (and life chances associated with) particular identities is contextually and ecologically contingent. We propose that an exciting and novel extension would involve integrating intersectionality with multilevel analyses of the contexts within which intersectional identities exist (e.g. schools, neighbourhoods, states).

Another innovative approach to using intersectionality theory in epidemiology involves Structural Equation Modelling (SEM). Measuring racism or sexism purely through racial groups or gender may be far too simple to detect the underlying power structures they seek to measure. SEMs would allow issues such as racism or sexism to be specified as unobservable latent variables. These latent constructs could then be measured using multiple variables to account for their multiple causes/dimensions. A further strength would be that causal pathways could also be explicitly specified and tested, allowing for the evaluation of particular pathways (or theories) through which inequalities are produced.

**Conclusion**

There is growing interest in applying intersectional approaches in health inequalities research. While theoretically rich, intersectionality has yet to be sufficiently tested in the health domain and there is a need for greater evidence and debate for how it can contribute to current population health approaches. We call for greater empirical evidence and discussion of the contribution of intersectional approaches for developing current approaches within population health research. By focusing attention on the upstream determinants of health, whilst also accounting for both between and within group differences, intersectional approaches help explore the extent to which interlocking systems of oppression and privilege become embodied as health disparities. We have outlined several key conceptual issues and methodological opportunities which we hope will help to drive the field towards a truly eco-epidemiological approach.

**References**

Capewell, S, Graham, H. 2010. Will Cardiovascular Disease Prevention Widen Health Inequalities? *PLoS Medicine* 7: e1000320.

Collins, PH. 1990. *Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment*. New York: Routledge.

Crenshaw, K. 1989. Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory, and Antiracist Politics. *University of Chicago Legal Forum*.

Damen, JAAG, Hooft, L, Schuit, E, *et al.* 2016. Prediction models for cardiovascular disease risk in the general population: systematic review. *BMJ* 353: i2416.

Davey Smith, G. 2011. Epidemiology, epigenetics and the ‘Gloomy Prospect’: Embracing randomness in population health research and practice. *International Journal of Epidemiology* 40: 537-562.

Doll, R, Peto, R, Wheatley, K, Gray, R, Sutherland, I. 1994. Mortality in relation to smoking: 40 years' observations on male British doctors. *BMJ* 309: 901.

Evans, CR. 2015. Innovative Approaches to Investigating Social Determinants of Health - Social Networks, Environmental Effects and Intersectionality. Doctoral dissertation, Harvard T.H. Chan School of Public Health. Accessed at: <https://dash.harvard.edu/handle/1/23205168>.

Green, MA, Strong, M, Razak, F, Subramanian, SV, Relton, C, Bissell, P. 2016. Who are the obese? A cluster analysis exploring subgroups of the obese. *Journal of Public Health* 38: 258-264.

Jones, K, Johnston, R, and Manley, D. 2016. Uncovering Interactions in Multivariate Contingency Tables: A Multi-Level Modelling Exploratory Approach. *Methodological Innovations* 9:1-17.

Manley, D, Johnston, R, Jones, K, Owen, D. 2015. Macro-, Meso- and Microscale Segregation: Modeling Changing Ethnic Residential Patterns in Auckland, New Zealand, 2001–2013. *Annals of the Association of American Geographers* 105: 951-967.

McCall, Leslie. 2005. The Complexity of Intersectionality. *Signs* 30(3):1771-800.

Rose, GA. 1985. Sick Individuals and Sick Populations, International Journal of Epidemiology. 14: 32-38.

Rose, GA. 1992. *The Strategy of Preventive Medicine*. Oxford, England: Oxford University Press.

Susser, M, Susser, E. 1996a. Choosing a Future for Epidemiology: I. Eras and Paradigms. *American Journal of Public Health* 86: 668-673.

Susser, M, Susser, E. 1996b. Choosing a future for epidemiology: II. From black box to Chinese boxes and eco-epidemiology. *American Journal of Public Health* 86: 674-677.

Wemrell, M, Mulinari, S, Merlo, J. 2017. Intersectionality and risk for ischemic heart disease in Sweden: categorical and anti-categorical approaches. *Social Science & Medicine* 177: 213-222.