25 Years of Multilevel Models in Social Epidemiology

A Story in Three Acts

Sam Harper

2023-06-14

Thanks for the invitation to discuss the recent history of multilevel modeling in social epidemiology. I'd like to do a bit of setting the scene, as my co-presenters are here to provide a lot more of the substantive and methodological discussion. So, a bit of background.

Let's go back to the early 1990s, filled with baggy clothes and questionable mash-ups of different musical genres.

One of the first papers to explicitly advocate for multilevel analysis in epidemiology was from Von Korff et al. in 1992. Noting the recent development and application of multilevel statistical models in the 1980s in demography, statistics, and education, they made a pitch for epidemiology to take up and use multilevel models, taking care to note that such an exercise depended on having explicit theories for why and how extra-individual factors may affect individual health. It's also worth noting that their main focus was on the ability of multilevel models to allow for the simultaneous estimating of both individual and contextual factors.

At the same time, a strong set of critiques of the discipline were coming from social epidemiologists, notably Krieger, Schwartz, Pearce and Ezra and Mervyn Susser, arguing that epidemiology's obsession with individual risk factors and failed behavioral trials were causing us to lose the population perspective that was crucial to past successes.

These critiques found additional support in the conceptual models that were also emerging at this time. Virtually all of these conceptual models for understanding the determinants of health placed some primacy on two things. One, that the traditional, individual-centered focus on epidemiology on risk factors for disease was inadequate for explaining population health and health inequalities. Two, that the process of generating distributions of health and disease involved multiple levels, including communities, neighborhoods, municipalities, states, etc. The development, adoption, and use of multilevel statistical models held enormous promise for highlighting and demonstrating the relevance of these new conceptual models, with the aim to bring back the population perspective.

So, at this time social epidemiology was in some respects primed for adoption of multilevel thinking given the long history of interest in place-based comparative health assessments in the tradition of the discipline's founders, the idea of communities as dynamic laboratories and, as previously noted, a strong emphasis on the population perspective as a tool to oppose ethe xcessive individualism of biomedicine.

With that background in mind, let's turn the Act 1 of our short story: The Big Idea. I would argue that there were really two big ideas that helped spur the adoption of multilevel statistical models in social epi. The first was the idea of simultaneous analysis of both individual and contextual factors. Prior to the development of the modern causal inference perspective now prevalent in most training programs, finding "independent" effects via multivariable regression was the goal of many etiologic studies. Early work from Pat O'Campo, Ana Diez-Roux, and George Kaplan in the US (though there were also a number of early studies in Europe as well) mainly focused on mutual adjustment and the search for 'independent' ecological or contextual exposures after adjustment for individual-level risk factors.

And it's hard to overstate the importance to social epidemiology of the early 1997 Science paper by the sociologist Robert Sampson and colleagues reporting that individual-level violent crime outcomes in Chicago were linked to neighborhood level measures of collective efficacy and informal social control. It linked a strong theoretical foundation with detailed field measurements, sophisticated multilevel models, and consideration of mechanisms.

Credit for early adoption and popularization of these methods should also be given to the 1998 paper by Diez-Roux that made a more specific case for adopting multilevel models. It remains a classic paper that laid out core conceptual and methodological issues related to defining contextual exposures, measurement challenges, and models. The majority of the focus remained, however, on the idea of simultaneous statistical adjustment. But she also made it clear that multilevel models were one way to help restore the population perspective to epidemiologic research.

The second 'Big Idea' that increased adoption of multilevel models was the issue of income inequality and health. Early ecological studies found that in rich countries population health was more strongly linked to income distribution than average income. By definition, inequality was a contextual measure, since it isn't a property of individuals. This idea prompted a large literature of multilevel investigations on how inequality might affect health, via relative comparisons, psychosocial effects, social capital, or other channels, a reflection of ideas present in the Science paper by Sampson et al.

Once the statistical issues had been more-or-less well described, much of the methodological discussions in the social epi literature focused on issues related to how to define neighborhoods, boundaries, how to develop, capture, or measure the relevant dimensions of context. As well as issues related whether and how to control for relevant individual-level variables.

By 2001 a systematic review of 25 neighborhood studies found fewer than half actually used multilevel studies, but continued to emphasize that nearly all found some statistically signifi-

cant associations after controlling for individual factors. Moreover, it maintained the idea that interventions at the contextual level were likely to be useful.

Social epidemiology largely continued down this path, with a series of multilevel tutorials from Juan Merlo that made a case that clustering and decomposing variance into between and within contexts, easily accomplished by random effects models, was a key to helping epidemiology make better decisions about interventions.

This idea had strong pull and by the early 2000s, researchers and public health advocates were making strong arguments that the time for intervening in neighborhoods had come.

Which brings us to Act 2, which I'll call A Crisis of Confidence.

The US Department of Housing and Urban Development had taken seriously the idea of intervening to change individuals' exposure to disadvantaged neighborhoods, designing and implementing an expensive trial called 'Moving to Opportunity' that randomized more than 4500 poor families to a housing voucher that mandated moving from higher to lower-poverty neighborhoods. The interim results of that trial were published in 2003. The intervention did lead to unconfounded differences in exposure to high-poverty neighborhoods, but the results were decidedly mixed: pretty clearly null effects for mothers on economic outcomes, but some positive and even some negative outcomes for kids, depending on gender and outcome. Like all trials, this one had many limitations, but the lack of transformational outcomes provided neighborhood effects skeptics with some ammunition.

Every story needs a villan (or a hero, depending on your perspective), and in 2004 the social epidemiologist Michael Oakes took up that role by launching a broadside against the neighborhood effects empirical literature, which generated an serious exchange with both Ana Diez-Roux and SV Subramanian, two leading multilevel modeling researchers.

Oakes effectively argued that multilevel modelers were too focused on the uncritical use of mutual adjustment, significance testing and random error, despite the fact that these were not the real problems to be overcome. What were the problems? The inability to tease apart individual and social exposures, endogeneity of neighborhood exposures, model extrapolation and spillover effects.

Meanwhile, the income inequality hypothesis was also having a crisis. Several multilevel papers found limited evidence that inequality was 'independently' associated with health, and the literature was highly heterogenous. Our systematic review in 2004 argued that nearly 100 multilevel studies had produced, at best, weak evidence.

Some of these discrepancies were do to explicit modeling choices. A testy exchange in Health Services Research between the economists Mellor and Milyo and Subramanian, Blakely, and Kawachi came to very different conclusions as to the value of whether a fixed or random effects approach was most appropriate for answering the question, with the fixed effects approach resulting in null evidence and random effects showing inequality was correlated with worse health.

By the late 2000s the neighborhood effects literature had matured, but some cracks were evident. Among 86 papers reviewed in 2007 nearly 80% were still using cross-sectional designs and mostly random effects, and among these the findings were inconsistent. And yet, the *idea* of the promise of multilevel models for revealing how to intervene on contexts remained strong.

Finally, our villain/hero Oakes had effectively given up on the whole project, arguing that we were just not learning much about causal effects of neighborhoods because multilvel studies simply refused to engage critically with the issue of causal identification.

Which brings us now to Act 3: is there 'A Way Forward'?

Recent years have actually seen a very productive merging of the huge increase in work on formal causal inference and work on multilevel models in social epidemiology. In particular, there has been much more attention focused on the key challenges to identification raised by Oakes and others. This has taken the form of stronger study designs like cluster RCTs and quasi-experiments. Moreover, an emphasis on utilizing longitudinal data to take advantage of individuals moving in and out of neighborhoods, as well as neighborhood changes themselves have tightened the focus on assumptions needed for causal inference. Finally, work on weighting methods to handle time-varying exposures and confounding have helped recent studies to deal with the problem of 'overcontrolling' for individual-level factors noted in prior work. All of these developments fit quite well within the framework of multilevel social epidemiology, though I admit are still not the most common approaches

In the early 2010s work by Hong and Raudenbush in education helped to push the methods forward for multilevel time-varying designs, and assumptions for inference were helped by directed acyclic graphs such as this one from Ari Nandi and Ichiro Kawachi. Researchers like Magda Cerda and Maria Glymour were early adopters of this improved work, and there continued to be substantive debates around modeling specifications, again with some controversy regarding the utility (or futility) of random effects models.

The work on causal inference also helped to stimulate a productive discussion regarding the disappointing MTO results. This focused largely on whether the ITT estimates from MTO were adequately answering the right question, or capturing the impact of duration of exposure to disadvantaged neighborhood conditions.

Using a novel design, Sampson, Sharkey and Raudenbush used marginal structural models to demonstrate how one could assess the impact of different durations of exposure to high-poverty neighborhoods, and found stronger effects when incorporating lagged impacts.

Although the positive merging of causal inference and multilevel models pushed the field forward, many studies using these designs found weaker evidence for neighborhood effects on health. Galston and Sharkey reviewed studies with credible causal designs in 2017 and noted decidedly mixed evidence, as well as examples demonstrating problems of self selection and confounding in studies with weaker designs. Nevertheless, they found the evidence was

heterogeneous and required keen attention to issues of the context, specific exposures and outcomes for neighborhood studies.

Finally, a word about cluster RCTs. One of the main challenges to the MTO trial was the fact that it changed where individuals lived rather than changing neighborhood conditions. Cluster-level interventions instead focus on changes places, perhaps more in keeping with the original spirit of contextual studies in epidemiology. Recent trials by Branas and colleagues intervening to remediate blighted housing lots in Philadelphia have shown both meaningful and cost-effective impacts on violent crime.

So, where does that leave us? On the whole, the embrace and adoption of multilevel models in social epidemiology has helped to spur innovation in the field: conceptually, methodologically and substantively. Recent work is taking causal inference seriously; however, one might say that the field still has yet to deliver on the big promises of the early papers advocating for multilevel models. Moreover, I would still say the modal neighborhood study remains cross sectional, a design unlikely to help at this point. Regarding the neighborhood effects literature, it seems at this point there is no overarching story, though we do have some strong evidence from high quality designs. This seems less true for adult health than for children, so incorporating a lifecourse perspective regarding duration of exposure is also likely to be fruitful. Finally, although resource intensive and complex, cluster-randomized designs are probably underutilized given the inferential strengths they bring to the table. Thank you.