Neighbourhood effects and the real world beyond randomized community trials: a reply to Michael J Oakes
From JUAN MERLO¹,² and BASILE CHAIX¹,³

We appreciate the positive opinion of Michael J Oakes¹ on our article² and his recognition of the contribution of our research group to the field of contextual epidemiology. We also understand that Oakes’ critics are not specifically directed at our publication but express a general concern on the use of observational approaches in research on neighbourhood and health.

In the absence of randomization the possibilities for confounding are theoretically enormous.³ Recognizing this may lead to a paralysing nihilism or to a radical ‘trialism’ (overemphasis of the advantages of randomized trials).⁴,⁵ Conversely, rather than rejecting observational studies, we believe it is our effort to reduce bias and confounding that constitutes a condition for the progress of epidemiology. Randomized trials, in spite of their theoretical suitability for causal inference, do not offer a solution to many epidemiological questions.³

Severe socioeconomic stratification, inferential support, and structural bias in observational multilevel analysis

Oakes argues that the individual characteristics leading to the ‘selection’ of individuals into a specific neighbourhood are entirely different from those leading to the selection of individuals into other contrasted neighbourhoods. This complete separation would produce a situation of structural confounding impossible to eradicate: specifying our individual-level regression model to get unconfounded neighbourhood variance, we would simultaneously eliminate all neighbourhood variance, leaving no variability to explain with neighbourhood variables.⁶

Oakes’ scenario of absolute separation, as a good example of ‘paralysing nihilism’, is theoretically possible but unrealistic. Whether socioeconomic stratification produces such an absolute confounding is an empirical question to address in each study.⁷,⁸ In our database, 34% of residents of deprived areas had a low income, as compared with 19% of residents of affluent areas. Therefore, our data do not support Oakes’ picture of absolute separation of individual attributes in contrasted neighbourhoods. Also, we were able to separate the population density and area socioeconomic effects, since the distribution of individuals across cells of combined area indicators (Table 1) was much less unbalanced than that approximated by Oakes.¹ In any case, such estimation problems can be mitigated by the use of large databases as the one we employed and Markov chain Monte Carlo estimation approaches.⁹

Oakes argues⁶ that multilevel analysis is suitable for investigating school effects on pupil performance but not for neighbourhood effects. The reason would be that teachers constitute exogenous forces influencing the pupils, in opposition to most neighbourhood effects that would be endogenous (i.e. determined by the neighbourhood composition in terms of various individual characteristics). We do not agree with Oakes’ argument. Indeed, many neighbourhood factors are exogenous forces influencing the residents (e.g. healthcare and sport facilities, urban design, etc.). Moreover, even endogenous neighbourhood effects can be properly investigated as long as social stratification is not as strong as in the extreme case considered by Oakes.

The reason for analysing within-neighbourhood clustering

Measures of neighbourhood variance complement information obtained from classical measures of association.¹⁰ Multilevel and spatial measures of variance are useful to assess whether health phenomena have a contextual dimension and the geographical scale of variations.¹⁰-¹⁶ We consider these indicators as descriptive but interesting measures of geographic heterogeneity but have never proposed such tools as a way to perform causal inference on ‘pure’ neighbourhood effects as Oakes seems to imply.

We see within-neighbourhood clustering of health outcomes as the result of a mixture of processes including selective migration driven by external forces of segregation and residential choices, integral external forces (e.g. urban design, healthcare facilities, air pollution) and contextual endogenous effects emerging from social interactions (e.g. sense of community, collective efficacy). Advances in the identification of these processes can only be expected by combining quantitative epidemiological approaches¹⁰-¹⁹ with hermeneutic qualitative methods and social observation.²⁰

Are randomized community trials the alternative to observational studies?

We are less confident than Oakes that randomized intervention studies constitute the canon for understanding neighbourhood effects. Sorensen have noted that ‘the randomized controlled design is the widely accepted paradigm for assessing
the effects of community interventions. However, this does not imply as Oakes infers that ‘the randomized community trial is canonical design for neighbourhood effect studies’. A community intervention is an exogenous force suitable for randomization and trial evaluation. Conversely, many neighbourhood effects are generated by the internal dynamics of the neighbourhoods and must be investigated as such since they cannot be experimentally recreated.

First, for reasons of practicability and costs, randomized community trials often include few communities/neighbourhoods, precluding generalizability and limiting the benefits of randomization. Second, owing to feasibility and ethical considerations, many hypotheses on neighbourhood effects cannot be tested through randomized trials. Third, randomized community interventions may lead to inferential problems when multiple causal pathways are involved, which is rather common because of the multifaceted nature of interventions. In this case, the only inference to perform is on the combined effects of the various action strategies implemented, which does not bring specific knowledge on a given neighbourhood effect outside an intervention context. In fact, causal inference on everyday-world neighbourhood effects seems in contradiction with the objective of community trials: whereas causal inference on out-of-intervention-context neighbourhood effects would need one to neutralize social placebo effects (changes induced in individuals by simply being the focus of attention), intervention planners understandably may wish to maximize it by involving as much as possible the whole community.

Finally, Oakes’ point of view that the non-exchangeability of individuals between neighbourhoods must be solved through randomization leads to a dramatic simplification of our tasks. In our view, the non-exchangeability of individuals between neighbourhoods is due to differences in individual resources and vulnerabilities that may confound, but also mediate or interact with neighbourhood effects, something of direct relevance to neighbourhood research rather than a nuisance to dissolve by randomization.

As a final point, we share the general concerns of Oakes regarding good observational epidemiological practice, including having strong a priori theories and hypotheses, and evaluating the consistency of the findings through different modelling strategies. Oakes’ insight is much appreciated and we, and surely our colleagues around the world, will consider his comments in future observational multilevel analyses.

### Table 1: Exact number of individuals in the sample per combination of categories of area indicators (area socioeconomic position and population density), females/males aged 55 years at baseline

<table>
<thead>
<tr>
<th>Area socioeconomic position</th>
<th>1st quartile</th>
<th>2nd quartile</th>
<th>3rd quartile</th>
<th>4th quartile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population density</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1st quartile</td>
<td>8954/8686</td>
<td>1470/1933</td>
<td>1038/859</td>
<td>3107/2385</td>
</tr>
<tr>
<td>2nd quartile</td>
<td>2973/3001</td>
<td>1936/1825</td>
<td>4437/3497</td>
<td>5428/5663</td>
</tr>
<tr>
<td>3rd quartile</td>
<td>1893/1728</td>
<td>3775/4131</td>
<td>4305/3886</td>
<td>4661/4200</td>
</tr>
<tr>
<td>4th quartile</td>
<td>835/510</td>
<td>7474/6048</td>
<td>4915/5666</td>
<td>1424/1692</td>
</tr>
</tbody>
</table>

\( a \) Our approach is to categorize continuous area variables with quartiles defined in the sample of individuals rather than in the sample of areas, in order to obtain a better comparability of the strength of associations between studies and because what is considered is an individual exposure to contextual circumstances. Because the population density factor is related to the number of individuals per area, the number of cases per quartile category exhibits a more unbalanced pattern than that of the number of individuals, as noted by Oakes.


LETTER TO THE EDITOR  1363

Genomics, epidemiology, and common complex diseases: let’s not throw out the baby with the bathwater!

From MUIN J KHOURY and MARTA GWINN

As public health professionals working to translate advances of genome-based research into population health benefits,1,2 we found the article by Buchanan et al.3 and associated commentaries4–10 fascinating and informative. We too are sceptical of ‘genohype’ and we are critical of the specious paradigm that leads directly from gene discovery to test development.11 We advocate an evidence-based approach that integrates knowledge from diverse fields—including genetic epidemiology3—to assess the clinical utility of genetic information for the benefit of population health.13

Therefore, we were astonished and disappointed to encounter Buchanan’s summary dismissal of genetic epidemiology as a misguided and hopeless quest for the philosopher’s stone. Delivered scarcely 3 years after completion of the Human Genome Project, this judgement is clearly premature. Population-based epidemiological research that makes the most of newly available information and techniques is just beginning. We should expect this research to take time—years, if not decades—to appropriately conduct, analyse, report, and synthesize.

Although we agree that public health programmes should continue to promote a healthy diet, adequate physical activity, and smoking cessation, it makes no sense to assert that ‘the preponderance of cases of complex chronic disease are owing to exogenous experience rather than endogenous genetic susceptibility’.10 Obviously, ‘people are not born with complex, late-onset disease’; on the other hand, people who continue to eat too much, spend too much time on the couch and smoke (despite vigorous public health education campaigns) do not always develop heart disease. Clearly, there is much more to learn about gene–environment interactions underlying these diseases and to use this knowledge in intervention efforts.

As we have argued elsewhere,14 the public health significance of genomic research on common complex diseases with strong environmental determinants lies not in finding new genetic ‘causes’ of these diseases but in helping us to better recognize and modify interacting environmental risk factors. Each investigation that increases our understanding of gene–environment interaction, etiological heterogeneity, pathogenesis, and natural history of common diseases adds to a knowledge base for estimating risks and guiding interventions to improve population health. Epidemiology is unique in offering a set of evolving tools and methods that are explicitly designed to observe disease variation in populations and reveal the joint effects of individual biology and behaviour in the context of social and physical environment. In an already complex world, human genetic variation is another dimension that is just now opening for exploration.15 For epidemiologists to retreat now would be to abandon the field just when they are needed most.

The concerns enumerated by Buchanan et al.3—including phenotypic and genotypic heterogeneity, the interplay between individual and ecological variables, the dynamic nature of environmental risk, chance, and bias—are all important and well-recognized challenges in epidemiological research. Nevertheless, we take issue with their assessment that ‘the lack of an obvious alternative does not justify continuing to invest in what does not work’. Indeed, there is no obvious alternative to epidemiology for translating genetic information from basic science to population health benefits but the assertion that it

National Office of Public Health Genomics, Centers for Disease Control and Prevention, 4770 Buford Highway, Atlanta, GA, USA.

*Corresponding author: E-mail: mkhoury@cdc.gov