AN INTEGRATED EPIDEMIOLOGICAL DESIGN TO IMPROVE MULTILEVEL ANALYSIS

José Miguel Martínez[1,2],Yutaka Yasui[3], Joan Benach[1,2], Fernando G. Benavides[1,2], Carles Muntaner[4], Ramón Clèries[5,6], Oscar Zurriaga[7], Miguel Angel Martínez-Beneito[7]

[1]Occupational Health Research Unit, Deparment of Experimental Health Sciences,Universitat Pompeu Fabra, Spain.
[2]CIBER de Epidemiología y Salud Pública (CIBERESP), Spain.
[3]Department of Public Health Sciences, School of Public Health, University of Alberta. Canada.
[4]Social Equity and Health Section, Centre for Addiction and Mental Health, University of Toronto. Canada.
[5]Servei d'Epidemiologia i Registre del Càncer de l'Institut Català d'Oncologia, IDIBELL, Spain.
[6]Departament Economia de l'Empresa de l'Escola Universitària d'Estudis Empresarials, Universitat Autònoma de Barcelona, Spain.
[7] Área de Epidemiología, Dirección General de Salud Pública, Generalitat Valenciana, Spain.

Multilevel analysis has been widely used to allow the simultaneous examination of the effects of individual- and group-level variables on individual health outcomes. In spite of its utility, multilevel design can have some drawbacks in the estimation of the risk factor effects when within-group variation of interest is small relative to between-group variation. An extreme case of this is a group-level risk factor which by definition has no within-group variation. Recently, an integrated design with a population-based estimating equation (PBEE) approach as its analysis method was proposed, where an aggregated health outcome of each group is utilized with an individual health outcome of sampled individuals in each group. In this paper, we describe the integrated epidemiological design and its advantages and disadvantages compared to the multilevel design. Specifically, using data simulated under different scenarios we show that integrated design has more power than multilevel in the estimation of individual- and group-level variables' effects on health outcomes when the between-group variability is higher than the within-group variability. In addition, the two designs were applied to a real-life data example of mortality following chronic kidney disease, illustrating actual magnitude of differences that might be encountered in practice.