

Dr Sandy Burden, CSSM

Title: Zones Issues for Small Area Health Data

Abstract: It is well established that in many cases the analysis of aggregate health data results in biased parameter estimates. However, the implications of modifying the geographic analysis zones used to aggregate and analyse the data have not been widely considered. In this study the effects of zoning on parameter estimates are investigated empirically at several scales of analysis. For an area level statistical model, zoning distributions are created for parameter estimates and statistics such as the population mean and regression coefficients. The zoning distribution defines the probability distribution or density function of the statistic over all possible sets of M zones that could be formed, given the constraints used in constructing the zones.

The analysis is undertaken using detailed population information. Realistic individual level data are simulated by combining unit record files from the Australian 2007-2008 National Health Survey with area level constraints from the 2006 Australian Census using combinatorial optimisation techniques. Empirical zoning distributions are created from the parameter estimates obtained from analysing the data summaries for each set of zones. The simulation of individual level data within small areas allows the implications of using aggregate data and a particular set of zones to be assessed for both a continuous and a binary response variable.