Assessing socioeconomic effects on different sized populations: To weight or not to weight?

N Frohlich, K C Carriere, L Potvin, C Black

Abstract

Objective—Researchers in health care often use ecological data from population aggregates of different sizes. This paper deals with a fundamental methodological issue relating to the use of such data. This study investigates the question of whether, in doing analyses involving different areas, the estimating equations should be weighted by the populations of those areas. It is argued that the correct answer to that question turns on some deep epistemological issues that have been little considered in the public health literature. Design—To illustrate the issue, an example is presented that estimates entitlements to primary physician visits in Manitoba, Canada based on age/gender and socioeconomic status using both population weighted and unweighted regression analyses.

Setting and subjects—The entire population of the province furnish the data. Primary care visits to physicians based on administrative data, demographics and a measure of socioeconomic status (SERI), based on census data, constitute the measures.

Results—Significant differences between weighted and unweighted analyses are shown to emerge, with the weighted analyses biasing entitlements towards the more populous and advantaged populations.

Conclusions—The authors endorse the position that, in certain problems, data analyses involving population aggregates unweighted by population size are more appropriate and normatively justifiable than are analyses weighted by population. In particular, when the aggregated units make sense, theoretically, as units, it is more appropriate to carry out the analyses without weighting by the size of the units. Unweighted analyses yield more valid estimations.

This paper deals with the use of ecological data from population aggregates of different sizes. It presents an issue that must be addressed when attempting to find a relation between various characteristics of population aggregates and some outcome variable. For example, if one were estimating utilisation of health care resources in various areas as a function of socioeconomic and demographic variables, would it be appropriate to weight the observations by the areas’ populations?

We argue that the correct answer to that question turns on some deep epistemological issues that have been little considered in the public health literature. We endorse the position that data analyses unweighted by population size are often more appropriate and normatively justifiable than unweighted analyses. In particular, when the units of analysis are aggregated individuals whose aggregations make sense as units, it is more appropriate to perform unweighted analyses. Unweighted analyses can yield less biased estimations of the aggregate level parameters. Alternatively, weighting might provide approximate solutions if one were interested in making inferences at the individual level from the aggregated data.

We use an example from Manitoba, Canada to demonstrate the effects of the methodological decision of weighting areas by their populations and show how that decision can lead to very different estimations of utilisation and entitlement.

A problem implicit in the traditional approach to ecological analyses

Rothman\(^2\) defines ecological studies as those using groups of people or areas as observation units rather than individuals. His recommendation is to weight each area or group by the size of the population because “the amount of information in each observational unit will differ.”(page 304).\(^2\) This prescription has been followed by some researchers who have linked resource allocation to the socioeconomic and demographic characteristics of the population.\(^3\)

By contrast, a number of empirical studies that address the role of social cohesion and social capital on health\(^1\) are unweighted correlational analyses of aggregate data (for example Kennedy et al\(^4\) and Kaplan et al\(^5\)).

There are a few reasons why weighting may have appealed to analysts.\(^6\) One is the traditional argument that risk factors are assumed to operate at the individual level and so the number of individuals with particular characteristics is assumed to be a relevant factor. Another reason is that many models are based on samples from larger populations and it is important that the model reflects the relevant characteristics of the underlying distribution of the larger population. This is particularly important when sampling is not proportional to the size of areas involved and the aim of the analysis is to estimate an “overall” mean rather than to estimate means for each group or area. Despite the substantive importance of this analytic choice, relatively little attention has been paid to the theoretical justification for this way of proceeding.

Implicit in the first argument is the notion that the question to be addressed is meaningful.
only at the level of the individual. The assumption is that were data available at the individual level the analysis would be carried out at that level; the resort to aggregate measures is only a function of the unavailability of individual data. There is a concern that the inferences made from aggregate level data, when applied to individual level phenomena, are vulnerable to the cross level biases that occur when confounding or effect modification is heterogeneous across analytical units. Such cross level biases result in a failure of the ecological effect to reflect the biological effect at the individual level and has been labelled the “ecological fallacy”.

There exist several means by which one can identify the presence of cross level bias and correct for the ecological fallacy.

On the other hand, critics have emerged who have begun to question the appropriateness of focusing solely on variables and phenomena at the individual level for an understanding of health and health outcomes. For example, Shy argues that epidemiology has become a biomedical discipline focused on the distribution and determinants of disease in groups of individuals and has disregarded the ecology of human health. Some have introduced aggregate level concepts such as collective lifestyle to render the idea that health and its determinants are shaped by collective attributes. Others have identified the concept of an “individualistic fallacy”. This occurs when population patterns of outcomes are erroneously presumed to be explicable only by individual level characteristics.

Another argument for the aptness of weighting by population turns on representativeness. That argument is applicable if the analysis is based on a sample from a larger population and the objective is primarily to estimate the overall population. When the analysis is based on an entire population, that argument loses much of its force. That caveat is particularly germane as the population is broken into areas that are believed to have unique area level characteristics that can affect the status of individuals and there is no reason to believe that those effects are a direct function of size. In those cases, an analysis unweighted by population would be more appropriate to measure the impact of those area level variables.

One possible means of dealing with these problems is the use of hierarchical modelling that separates individual and area effects. However, weighting versus unweighting remains an issue in that estimating procedure. Moreover, that form of modelling may not be appropriate for analyses that attempt to identify only area level effects. For example, Ross et al have recently produced a univariate analysis of the relation between income inequality and mortality in Canada and the United States based on a state and province level analysis. Their analysis evidences how the aggregated level inference. In performing an analysis, then, one must consider the arguments for and against using the population size of the area as a weight variable in a population based analysis. The immediate effect of weighting area level observations by their populations is simple to describe: areas with larger populations have a greater impact in the estimation of
To weight or not to weight?

Consider, for example, a suggestion made by Noralou Roos. Suppose the World Health Organisation were attempting to determine the link between the levels of industrialisation of nations and their health statuses. Suppose they were to use a population weighted model. China and India both have huge populations and although they are both at the lower end of the development scale, they would largely determine the result. The variation between these two countries would determine much of the nature of the relation even though it would be based on a very small spread in the independent variable (level of industrialisation). The effects of economic development and industrialisation at the higher end of the scale, represented in the European and North American countries (of smaller populations) would be largely masked as would the contribution of much of Africa and South America. Weighting by population would underestimate the effects of the smaller population countries and overestimate those of the larger countries. On the other hand an unweighted analysis runs the risk of having many small countries play a disproportionate part in determining the relation.

How does this example lead one to argue for an unweighted analysis as more appropriate? Firstly, theory and evidence indicate that the areas under consideration can be treated as entities: they have one or more characteristics that have effects on an outcome that are meaningful at the population level. Secondly, the variables of interest are not distributed in the same way in the less populated areas as they are in the more populated areas. It should be noted, that these conditions preclude the aggregation of smaller areas into larger areas to balance the sizes of the units under analysis, and thereby to finesse the weighting problem. And finally, there is a correlation between the variables of interest and the size of the population that is not deemed to be of theoretical importance. Taken together, these characteristics lead one to conclude that an unweighted analysis more accurately reflects the underlying relation between the explanatory variable and the outcome. A population weighted analysis overrepresents the larger areas and underrepresents the smaller areas and hence biases the estimate.

Methods

We present an example that uses socioeconomic and demographic data to derive a measure of entitlement to physician services. Data from the entire population of the Province of Manitoba were aggregated at the level of physician service area (PSAs). These areas were geographically contiguous towns or areas within municipalities that shared the bulk of their visits to a given concentration of physicians. These data were used to identify the relation between demographic and socioeconomic indicators (which were treated as proxies for need for health services) and utilisation of primary care visits to physicians.

This example is presented primarily for purposes of illustrating the methodological point we argue. However, it is substantively important in its own right because, increasingly, providers of comprehensive health care have begun to realise that the socioeconomic and demographic characteristics of their client populations can have a profound effect on the level of services demanded (and, possibly, warranted). In Canada, where the provinces fund these services centrally, there has been a move towards the establishment of regional health authorities who service diverse populations. A similar challenge faces HMOs in the United States operating multiple centres in areas with disparate populations.

The Province of Manitoba has recently established provisions for 12 regional health authorities. One of the first policy challenges has been determining the appropriate levels of regional funding. Central to that concern is the question of exactly how various indicators should enter into the determination of funding entitlements. Frohlich and Carriere conducted analyses to aid the development of needs based formulas for funding physician services in the province. Their analysis used socioeconomic and demographic indicators to estimate a reasonable basis for allocating resources. It is well established that need and utilisation patterns in population vary as an interactive function of age and gender. Consequently, it is standard practice, in comparing different population aggregates, to adjust for age and gender differences in populations.

Another factor, widely recognised as affecting health status, and hence probably need for services, is the socioeconomic status of the population. An index of socioeconomic status has been developed for the Manitoba population. The Socio-Economic Risk Index (SERI) is based on census data, which reflects the relation between the socioeconomic status of the population and health status. Using data from 282 municipalities, a broad swath of socioeconomic variables from the 1986 census, were modelled against five indicators of poor health and lifestyle, aggregated into an index. Six census variables were found to be closely related to this index and to explain a significant amount of the variance (59%). Three of these were positively related to poor health: the percentage of the labour force unemployed: ages 15 to 24, the percentage unemployed between 45 and 54, and the percentage of single parent female households. Three were negatively related: the percentage of the population between the ages of 25 and 34 having graduated high school, the percentage of females participating in the labour force and the average dwelling value. The predicted health statuses from the estimated equation were used to construct a standardised index: the SERI with a mean of 0 and a standard deviation of 1. The resulting index was shown to be strongly related to a number of other measures of health, such as premature mortality. In the analysis of primary care visit entitlement, that index was used as a measure of socioeconomic status.

The analysis was initiated on the premise that a first cut at entitlement to physician visits...
might be calculated by fitting age, gender and socioeconomic status (SERI) in a model to estimate average existing patterns of visits across PSAs. The calculated averages could then furnish the basis for determining an average entitlement for each area based on its age/gender mix and socioeconomic status.

The rationale for this as a first step is the premise that the universally funded medical system is somewhat, but imperfectly, responsive to the underlying health status of the population in the province. An estimate of average visits adjusted for age, gender, and socioeconomic status could serve as the basis for a first cut at estimating average entitlements to primary care visits for each area. For example, variations in the supply of physicians might well affect actual visit rates. Hence, areas with low levels of supply would probably show up as having visit rates that were lower than their estimated entitlements (given their age, gender, and SERI composition).

The model was constructed by identifying all ambulatory visits in the province outside of hospitals in 1993/94 at the individual level. Starting at the municipality level (259 municipalities) the province was divided into 58 PSAs. These areas were defined in terms of where individuals were getting the bulk of their primary care visits. The defining characteristic was a concentration of population and an associated set of physicians (usually no fewer than five) serving that population.

Using per capita visits as the dependent variable, a multiple linear model was constructed. The population in each PSA was divided into 21 age by gender categories: ages 0–1, 2–4, 5–9, 10–14, ..., 90–94, and 95+ years old by male or female. The outcome to be analysed was the actual rate of ambulatory visits to physicians. Visits to nursing stations were included. This reduces, somewhat, the issue of low supply of physicians in remote areas. Thus each PSA was represented by 42 data points. Each represented an average for a gender and age category and its associated average primary care visit level. In all, this comprised a total of 2436 observations. The areas varied in population size from 1020 to 120 250 and the age and gender groups varied in size from 1 to 6554.

The equation to be estimated was:

\[
\text{average visits per capita} = f(\text{age}|\text{gender}|\text{SERI})
\]

where "|" indicates the model incorporates all two and three way interaction effects.

That multiple linear analysis is really nothing more than a kind of averaging of utilisation patterns of ambulatory visits according to the areas’ age and gender groupings and SERI scores. The coefficient of each of the input variables in the model represents the impact of that characteristic on utilisation. It gives an estimate of the entitlement to visits for a member that category. Three characteristics (age, gender, and SERI of the PSA) and their interactions determine how many visits, on average, any individual is entitled to, given the prevailing pattern of visits in the province. The overall goodness of fit of the model shows the extent to which the areas in the province are receiving average levels of ambulatory visits in terms of their age/gender/SERI characteristics. The entitlement of an individual would be determined by the estimating equation’s coefficients.* One could use the individuals’ entitlements to calculate what the average expected ambulatory visit level would be for each area based solely on the age/gender/SERI characteristics of the individuals in that area.

This descriptive model could be used prescriptively. Each area’s relative entitlement would be defined by the mix of ages and genders in its population and their socioeconomic status as measured by the SERI. An area’s entitlement to visits would be the sum of the average visit rates in each age/gender category, suitably modified by the SERI factors, multiplied by the population size in that category. This entitlement is a sort of smoothing out—giving each area an expected allocation on the basis of its population’s age/gender/SERI distribution based on actual utilisation patterns.

The question of interest to us in this modelling exercise is whether to weight the estimating equation by the populations of the PSAs. Recall that each data point represents an age/gender category in a PSA. The actual number of individuals in any category in an area is the size of the population stratum in that area. One possibility is to perform a population weighted analysis taking into account the different sizes of the areas. That would be consistent with much current practice. Proceeding in that way, each data point (age/gender stratum in an area) would be weighted by the population in that stratum in that area. That would give weight to an area’s population in determining an average utilisation pattern. Hence, the patterns in more populous areas would play a greater part in defining an average pattern than would the patterns of more sparsely populated areas. The resulting “average” pattern would

---

* For example, the entitlement of a 4 year old girl from an area with a SERI of 1 would be calculated using the coefficients of the variables and their cross terms as follows:


www.jech.com
To weight or not to weight?

Visits/person entitlements for 1993/4.

Results

Figure 1 shows the different effects of population weighted and unweighted analyses. The figure contains two sets of 54 data points (each point corresponding to the PSAs per capita entitlement according to their age/gender/SERI characteristics, as described above). In figure 1 the entitlements are plotted against the directly adjusted premature mortality rates for individuals between the ages of 0 and 74 years of age (a measure of poor health status and hence indirectly of need for health services).

The estimated slopes of the two lines are very different. The lines in the unweighted analysis (through the triangles) shows a much steeper slope than in the weighted analysis (through the dots). The estimated slopes are 0.643 and 0.257 respectively. The unweighted analysis produces a stronger relation between entitlement and premature mortality than does the weighted analysis.

A complementary look at the differential effect of the two models can be obtained by plotting entitlements against the SERI. Figure 2 shows the relation between the SERI scores of areas and entitlements using weighted and unweighted analyses. Again the unweighted regression line is much steeper than the weighted line.

The differences in allocations under the alternative weighting systems are quite dramatic. Table 1 shows the different allocations to the Winnipeg and Northern (remote and poor) regions of the province under the two treatments, along with their observed crude visit rates/per capita, their SERI scores and directly adjusted premature mortality rates. PSAs are listed in order of decreasing SERI in each of the two blocs—that is, in order of decreasing socioeconomic risk.

From the table it is clear how the population weighted analysis allocates more per capita visits to the larger and better off Winnipeg PSAs (in terms of SERI and premature mortality) than does the unweighted analysis. The three relatively poor Winnipeg PSAs do worse under the weighted allocations than under the unweighted allocations. The same is

* There are other ways in which population size could be entered into the analysis. One could conduct a variance weighted analysis. We also did that, but the analysis lead to similar conclusions and so is not reported here. Alternatively, population size in itself could be considered an independent variable which either directly, or indirectly, affects the outcome variable. So, for example, if physician supply were expected to be greater or transportation costs lower, in areas of high population, then utilisation rates might be dependent on the size of the area.
Table 1  Allocations of ambulatory visits per capita population weighted and unweighted analyses compared

<table>
<thead>
<tr>
<th>Winnipeg PSAs</th>
<th>Population weighted per capita allocation</th>
<th>Population unweighted per capita allocation</th>
<th>Observed crude visits/capita</th>
<th>Premature mortality (0–74 y)</th>
<th>SERI scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Winnipeg Inner Core</td>
<td>5.073</td>
<td>7.552</td>
<td>6.25</td>
<td>7.206</td>
<td>0.855</td>
</tr>
<tr>
<td>Winnipeg Outer Core</td>
<td>5.417</td>
<td>6.164</td>
<td>5.44</td>
<td>4.344</td>
<td>0.225</td>
</tr>
<tr>
<td>Winnipeg Old St Bon</td>
<td>5.475</td>
<td>5.849</td>
<td>5.97</td>
<td>4.600</td>
<td>0.203</td>
</tr>
<tr>
<td>Winnipeg NE</td>
<td>4.643</td>
<td>4.281</td>
<td>4.69</td>
<td>3.100</td>
<td>-0.774</td>
</tr>
<tr>
<td>Winnipeg W</td>
<td>4.825</td>
<td>4.406</td>
<td>5.20</td>
<td>3.225</td>
<td>-0.795</td>
</tr>
<tr>
<td>Winnipeg NW</td>
<td>4.574</td>
<td>4.165</td>
<td>5.06</td>
<td>3.093</td>
<td>-0.823</td>
</tr>
<tr>
<td>Winnipeg S Central</td>
<td>4.805</td>
<td>4.152</td>
<td>5.41</td>
<td>3.050</td>
<td>-0.970</td>
</tr>
<tr>
<td>Winnipeg SE</td>
<td>4.236</td>
<td>3.556</td>
<td>4.73</td>
<td>2.905</td>
<td>-1.142</td>
</tr>
<tr>
<td>Winnipeg SW</td>
<td>4.204</td>
<td>3.265</td>
<td>4.64</td>
<td>2.735</td>
<td>-1.312</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Northern PSAs</th>
<th>Population weighted per capita allocation</th>
<th>Population unweighted per capita allocation</th>
<th>Observed crude visits/capita</th>
<th>Premature mortality (0–74 y)</th>
<th>SERI scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Island Lake</td>
<td>6.646</td>
<td>9.145</td>
<td>10.74</td>
<td>3.907</td>
<td>3.389</td>
</tr>
<tr>
<td>Norway Cross</td>
<td>6.470</td>
<td>8.744</td>
<td>7.33</td>
<td>5.229</td>
<td>2.758</td>
</tr>
<tr>
<td>Leaf Rapids</td>
<td>4.896</td>
<td>5.462</td>
<td>3.92</td>
<td>4.038</td>
<td>0.406</td>
</tr>
<tr>
<td>The Pas</td>
<td>5.095</td>
<td>5.710</td>
<td>4.55</td>
<td>5.459</td>
<td>0.284</td>
</tr>
<tr>
<td>Thompson</td>
<td>4.902</td>
<td>5.385</td>
<td>5.76</td>
<td>4.863</td>
<td>0.251</td>
</tr>
<tr>
<td>Flin Flon</td>
<td>5.169</td>
<td>5.757</td>
<td>5.25</td>
<td>4.971</td>
<td>0.132</td>
</tr>
<tr>
<td>Gillam</td>
<td>4.546</td>
<td>4.595</td>
<td>5.67</td>
<td>5.063</td>
<td>-0.328</td>
</tr>
<tr>
<td>Churchill</td>
<td>4.603</td>
<td>4.667</td>
<td>1.07</td>
<td>5.252</td>
<td>-0.336</td>
</tr>
</tbody>
</table>

true of all the Northern areas. A population weighted allocation gives all Northern regions lower per capita entitlements than does the unweighted analysis and the differences seem to be larger the worse off the population of the PSA. In the worst off areas of Island Lake, Oxford House and Norway Cross the differences are very great indeed: in the order of two visits per capita. These differences in allocations under weighted and unweighted analyses can be seen directly by computing the differences in entitlements (weighted analysis minus unweighted analysis) for each PSA and graphing it, again, in order of increasing SERI. Figure 3 presents that picture. Here the normative implications of choosing between the two estimations are clear. The five best off and more populous PSAs show a positive gain if weighted analysis is opted for over an unweighted analysis while the rest of the PSAs lose visits under that choice. And the worse off the PSA in terms of SERI, the more they lose. The relation is a very strong one. One can look at differences in entitlements between the weighted and unweighted estimates as they relate to the SERI of the PSA. A linear regression model estimating that relation is significant at the 0.001 level and explains over 95% of the variance.

Discussion

The differences noted above offer evidence of very different treatments of PSAs under the two analyses. The spread between the entitlements of the best and worst off PSAs is greater in the weighted as compared with the unweighted analysis. In the weighted analysis the range is (6.89–4.20) or 2.69 visits per capita while in the unweighted analysis it is (9.78–3.26) or 6.52 visits per capita.

The data reveal that the weighted analysis decreases the entitlements (from observed actual rates) of the better off Winnipeg areas only marginally (between 0.05 and 0.6 visits per capita), but it also decreases the entitlements of the poorer urban areas as well (0.02 to 0.5). In the North, it decreases the per capita entitlements of all save three of the Northern areas by even wider margins. By contrast the unweighted analysis reduces the entitlements of many of the better off Winnipeg areas quite substantially while increasing the entitlements of two of the three worse off Winnipeg areas. In the Northern areas the unweighted analysis increases the entitlements of two more areas than does the weighted analysis, but where reductions occur, they are generally much less severe.

So the general effect of an unweighted analysis is to smooth the entitlements in accordance with expectations and to favour the worse off areas more strongly.

Unweighted analysis shifts considerably more resources to the worse off areas. Which is the more appropriate allocation?

There is a straightforward argument that can be made in favour of the unweighted analysis in this case. Universal access to primary care visits can be assumed to be somewhat responsive to underlying need. There should be some rough but imperfect correspondence between the use of visits and measures of health status and/or other measures of need. A few disturbing factors might mitigate against a better fit between need and utilisation. One is the availability of physicians (both supply level and in terms of proximity), another is the availability of information regarding access to the system. Any calculation of entitlement should be expected to correct for disturbing factors and to reallocate per/capita entitlements away from the less needy towards the more needy areas. It should be expected to increase the variance between the worst off and best off areas in favour of the former. Of course, question of the appropriate coefficient for needs adjustment has bedevilled needs based planners. Exactly how many additional health care resources ought to be allocated to areas per unit increment in the measure (either premature mortality or SERI) is an issue that remains open but our analysis does allow us to draw some general conclusions.
Conclusions: To weight or not to weight

A purely theoretical argument can be made for the use of an analysis unweighted by population in a number of cases. When one performs an analysis weighted by population, then the population size of each area enters into the equation as a factor determining entitlement to health care resources. We know of no studies that indicate that the size the population of a community, in itself, affects the health status and need for health care services of the population. Hence, it is not appropriate to weight analyses by the population. However, as a referee has noted, other factors, highly correlated with population size, such as supply of services, or travel distance to access services, may affect utilisation. In those cases, care should be taken to identify any variables which are candidates to affect entitlement and they should be included in the model. If there is a suspicion that some variables, such as supply factors, are not measurable but are highly correlated with population size, then population size might be used, but only as a self standing factor.

There are, however, two important caveats to this argument. An unweighted analysis allows for each unit to represent a pattern of utilisation across the variables in question and to provide coefficients for utilisation across PSAs as the units of analysis. An unweighted analysis treats utilisation as a community (or PSA) level phenomenon. The SERI is not an individual measure but a measure that characterises background conditions in an area. Hence, the appropriate unit of analysis is not the individual, but the area. Another general conclusion is, therefore, that weighting by population is not appropriate in such cases.

There are, however, two important caveats to this argument. An unweighted analysis is subject to fluctuation (and hence manipulation) as a function of the way in which areas are defined. It is important that the areas chosen for analysis have some claim to theoretical relevance for the issue in question, and not be chosen arbitrarily. Moreover, assigning each community equal weight allows very small communities the same impact on the definition of average utilisation patterns as very large communities.

Nevertheless, the burden of the previous arguments lead us to conclude that population size should not be used as a weighting variable in an analysis when area effects are anticipated. In the example we have used, and by implication in all cases that share its characteristics, we know that the choice to weight or not can have major implications for the estimate obtained: its accuracy and its normative impact. The question of weighting by population in estimating relations, then, is a substantive and normative issue that needs careful consideration.

The authors would like to thank Noralou Roos, David Friesen and Leonard MacWilliam for their help in the initial project whence this argument sprung. We would also like to thank the community within the Manitoba Centre for Health Policy and Evaluation for innumerable and unspottable suggestions over numerous drafts of this argument. Three anonymous referees have also made a number of helpful suggestions. Funding support for the research was provided by the Centre. K C Carriveau was funded in part by Natural Sciences and Engineering Research Council of Canada, Alberta Heritage Foundation for Medical Research and National Health Research and Development Program. L Povin is funded by the Medical Research Council of Canada through a Scientist Career Award (MRC H3–17299–AP09720).

Conflicts of interest: none.


