Is there an association between shigellosis incidence and socioeconomic status in metropolitan Haifa?

Jonathan Dubnov, MD, MPH, Lisa Rubin, MD, MPH, and Shmuel Rishpon, MD, MPH
Haifa, Israel

Background: Shigellosis incidence rates in Israel have declined continuously over the past 50 years, but they remain 20 times greater than those in the United States. Socioeconomic factors may influence shigellosis morbidity, but this may be difficult to demonstrate in the absence of data for individual patients and when using composite rates for large geographic areas. Use of census tract data for small, relatively homogeneous geographic areas may lessen the effects of the “ecological fallacy.” The present study analyzes the effect of socioeconomic status (SES) on shigellosis morbidity in the Haifa metropolitan region.

Methods: The study population consisted of the 7 cities in the Haifa subdistrict that constitute the greater metropolitan region. Cases of shigellosis reported during the years 2000 and 2001 were mapped, and age-standardized rates were calculated for the census tract areas. The incidence rates were then compared with the SES category of the census tract using the Kruskal-Wallis test.

Results: No association was found between incidence rates of shigellosis and SES category of the census tract areas in the Haifa metropolitan area for the years 2000 and 2001 ($\chi^2 = 0.440; P = .803$).

Conclusion: We found no association between shigellosis morbidity and socioeconomic status. This finding is probably real and not the result of reporting bias. Analysis of morbidity using small geographical units such as census tracts is more accurate than analysis using large geographical areas such as cities. (Am J Infect Control 2004;32:274-7.)

Shigellosis is an acute gastrointestinal disease with fecal-oral transmission that can cause serious complications and occasionally death. More than 160 million cases are reported yearly worldwide, with 1.1 million deaths. Most reported cases occur in developing countries, but sporadic cases and outbreaks remain common even in developed countries despite the significant reduction in incidence that has occurred since the 1980s. Israel has seen a 50% decline in incidence rates over the past 15 years. The decline in morbidity occurred predominantly among the Jewish population and was accompanied by changes in prevalent strains. More than 90% of the reported cases in Israel are caused by Shigella sonnei. The number of cases caused by Shigella flexneri has declined 3-fold over the past 20 years. There has been a concomitant increase in the average age of infection. Between 1970 and 1990, the age-specific incidence rate among infants was greater than that of all other age groups in the Arab population. In the early 1990s, the age distribution of morbidity in the Arab population changed and became comparable to that seen in the Jewish population, with the age of maximal incidence being 1 to 4 years. Despite the consistent decline in morbidity over the last 50 years, the incidence of disease remains high and is 20 times that reported in the United States. These differences in morbidity between the Jewish and Arab populations may reflect differences in socioeconomic status in addition to differences in cultural and hygienic practices.

Socioeconomic factors have been reported to influence shigellosis morbidity, but these are difficult to demonstrate in the absence of these data for individual patients and when using rates calculated for large geographical areas. In the absence of individual-level data, composite-level data is frequently used. The assumption that the estimated average exposure of a particular geographical area is equivalent to the individual's actual exposure in that geographical area can cause the methodological error known as the “ecological fallacy.” The larger and the more heterogeneous the geographic area for which composite data is used, the greater the variance of exposure within that area is likely to be. The socioeconomic status (SES) index calculated for an entire city is the weighted average of its constituent neighborhoods or census tracts, which may have considerable variance among
them. Studies comparing morbidity between countries or even cities suffer from this flaw. Analyses using smaller geographical units such as census tracts, which are likely to have a more homogeneous population, may reduce this error. Such analyses are now possible using Health Geographical Information Systems. 

The present study analyzes morbidity from shigellosis in the Haifa metropolitan area and its possible association with socioeconomic status.

METHODS

Study population

The study population consisted of 7 cities in the Haifa subdistrict that constitute the greater Haifa metropolitan area: Qiryat Motzkin (City A), Qiryat Ata (City B), Qiryat Bialik (City C), Qiryat Yam (City D), Tirat Hacarmel (City E), Haifa (City F), and Nesher (City G). The National Bureau of Statistics (NBS) divides the area into 160 census tracts. The population of these cities during the years 2000 and 2001 numbered 465,000.

Incidence data

All cases of shigellosis reported by physicians and laboratories to the Haifa District Health Office as mandated by law were entered into a computerized registry of notifiable diseases. Only laboratory-confirmed cases were included. The registry utilized Microsoft Access software. Data included demographic variables (including address, diagnosis-coded according to ICD 9 CM) and clinical and laboratory information. The study population was limited to those persons whose address, as listed in the population registry of the Ministry of Interior, was within the geographical boundaries of the Haifa subdistrict. We did not analyze cases of shigellosis by Arab and Jewish population separately for two reasons: more than 94% of the study area is Jewish and only 4% of cases occurred in Arabs.

Accurate addresses that could be mapped were available for 296 of the 311 reported cases (95.2%); 4 cases of the 296 were excluded from further analysis because insufficient data prevented calculating age-standardized incidence rates for the census tracts in which they occurred. In all, 292 (93.9%) of the 311 reported cases were included in the final analysis.

Mapping

Individual cases were mapped using Arcview GIS. City maps were provided by the NBS and included the municipalities, census tracts, and streets.

Census tracts are small, discrete geographical areas that the NBS utilizes when collecting and analyzing census data. Because census tracts cover small areas, the population within a particular tract is fairly homogeneous. We used census tract data from the most recent (1995) census. The information available for most of the census tracts included distribution of sex, age, and religion of the population as well as the SES index. This index is a weighted average of 15 variables, including income, educational level, and housing density, calculated by NBS using factor analysis. In the study area, SES information was available for 141 (88.1%) of the 160 census tracts. The 19 census tracts for which there was no information are industrial zones that are very sparsely populated. We constructed maps depicting cases of shigellosis in the 7 cities for the years 2000 and 2001. Cases were reported in 103 (75.0%) of 141 census tracts of the study cities.

To examine the possibility of reporting bias, we compared the distribution of SES of census tracts with reported cases of shigellosis to that of census tracts without reported cases. No differences in the distribution of SES were found between tracts with and without shigellosis incidence. The age-standardized incidence rates (ASR) using direct standardization were calculated for 99 of the 103 census tracts. Four tracts were excluded from further analysis because insufficient data prevented calculating ASR.

Analysis

Analysis was performed using SPSS 10.0. The distribution of the SES index (a continuous variable) of the census tracts was examined and found to be not normal. The NBS assigns every census tract to one of 20 ranked clusters based on the calculated SES. We divided the census tracts examined into 3 categories based on these NBS-ranked clusters: low SES (clusters 1-8), middle SES (clusters 9-12), and high SES (clusters 13-20). The null hypothesis for this final stage of the analysis was that there would be no association between the incidence of shigellosis and the SES category of the census tract. This was examined using the Kruskal-Wallis test (equivalent to analysis of variance nonparametric test) using a continuous variable (incidence rate) compared with a categorical variable (SES category) and calculating the Kruskal-Wallis $\chi^2$. To exclude type II error, we performed a power calculation based on Cohen’s medium and high effect sizes. We found that the power was above 70% and supposed it to be adequate for testing of the null hypothesis. A value of $P < .05$ was considered statistically significant and sufficient to refute the null hypothesis.

RESULTS

During the 2-year study period, 311 cases of shigellosis were reported in the 7 cities. They constituted 85% of the total morbidity in the subdistrict.
whereas the study population constituted 89% of the subdistrict population. Of the 311 cases, 19 were excluded from further analysis because insufficient data prevented accurate mapping or calculating age-standardized incidence rates for census tracts. The number of cases of shigellosis included in further analysis was 292 (93.9%) of the 311 reported cases.

The average age-standardized incidence rate calculated for 99 census tracts was 64.0 per 100,000 population (SD 49.2). The lowest age-standardized rates were those of City A and City B, whereas the highest rate was that of City G ($P < .05$) (Table 1). No statistically significant differences were found among the 3 groups ($P = .803$). The null hypothesis was upheld, and no association was found between SES category and shigellosis morbidity.

**DISCUSSION**

In contrast to previous reports, we did not find differences in age-adjusted incidence of shigellosis in populations with differing SES. Theoretically our finding could be due to a lack of power (type II error), selective reporting, mapping bias, or the ecological bias.

In general, nonparametric tests are less powerful than parametric tests, so when nonsignificant results are found, type II error should be excluded. We performed a power calculation, which was above 70%, and supposed that to be adequate for testing of the null hypothesis. Calculated exact significance levels did not provide statistically significant differences among the 3 groups of SES, so a type II error probably is not relevant for our findings.

Selective bias in disease reporting is possible since the reports are based on routine reports of shigellosis isolates received from the clinical laboratories of all the sick funds and hospitals in the area. Review of annual epidemiological reports from the subdistrict did not reveal any disparities in the reporting between the sick fund laboratories or hospital laboratories, but it is possible that physicians from certain areas might be more or less likely to request stool cultures when presented with a patient with diarrhea.

As mentioned above, no differences in the distribution of SES were found between tracts with and without reported cases of shigellosis. This finding might indicate equitable distribution and availability of health services, regardless of social class, as a result of the implementation of the National Health Insurance Law in 1995.

Since we were able to successfully map 95% of the reported cases, a bias in mapping probably does not account for our findings.

As in all ecological studies, our findings could reflect the ecological fallacy, that is, the assumption that the average exposure (in our case, SES) in a geographical area reflects the actual exposure of the individual, an assumption that may lead to false conclusions. The

**Table 1. Age-adjusted incidence rates of shigellosis by census tracts in 7 cities of metropolitan Haifa**

<table>
<thead>
<tr>
<th>City</th>
<th>No. of census tracts with reported cases (total no. of tracts)</th>
<th>SD</th>
<th>Age-adjusted incidence rate</th>
<th>$P$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td>City A</td>
<td>8 (11)</td>
<td>21.1</td>
<td>39.4</td>
<td>Reference</td>
</tr>
<tr>
<td>City B</td>
<td>9 (12)</td>
<td>17.1</td>
<td>39.4</td>
<td>NS</td>
</tr>
<tr>
<td>City C</td>
<td>9 (11)</td>
<td>20.0</td>
<td>45.9</td>
<td>NS</td>
</tr>
<tr>
<td>City D</td>
<td>8 (13)</td>
<td>32.7</td>
<td>51.6</td>
<td>NS</td>
</tr>
<tr>
<td>City E</td>
<td>4 (7)</td>
<td>61.3</td>
<td>72.5</td>
<td>NS</td>
</tr>
<tr>
<td>City F</td>
<td>57 (101)</td>
<td>56.2</td>
<td>73.3</td>
<td>NS</td>
</tr>
<tr>
<td>City G</td>
<td>4 (5)</td>
<td>58.3</td>
<td>92.4 ($&lt; .05$)</td>
<td></td>
</tr>
</tbody>
</table>

SD, Standard deviation; NS, nonsignificant.
use of smaller geographical and population units of analysis has been proposed in order to reduce this source of bias.\textsuperscript{7,8} Census tracts are relatively small geostatistical units, which are more likely to be homogeneous in population attributes than are larger geographical localities such as cities or districts. In the present study, we used these units to reduce the effect of the ecological bias. As mentioned above, we did not analyze cases of shigellosis by Arab and Jewish populations separately because the study area is predominantly Jewish and few cases were reported among the relatively small Arab population.

Differences in shigellosis morbidity between the Jewish population and rural Arab and bedouin populations have been described,\textsuperscript{11,12} but no studies have been published that examine the effect of SES within an area with a predominantly Jewish population, so a simple comparison with previous reports is not possible. We believe that our finding may reflect the true absence of an association between SES and shigellosis morbidity in the Haifa metropolitan area. This finding may reflect comparable hygienic living conditions for populations of differing SES in the Haifa metropolitan area.

The low infective dose of shigellosis might also contribute to the absence of an association between SES and morbidity. Even if there were a difference in hygienic conditions between different social classes, with a pathogen of such high infectivity, such conditions may not affect incidence rates.

The majority of shigellosis cases occur in children. In Israel in general and in the Haifa area in particular, all social classes have high use rates of day care centers and preschool facilities. This fact might also account for the lack of an observed association. In public health, the cycle of poverty-illness-poverty is well recognized. The present study, in contrast to previous ones, does not document an association between social class and shigellosis morbidity. This is a reassuring finding, particularly in light of increasing socioeconomic gaps in Israel in recent years.

Analyzing morbidity data according to small, homogeneous geostatistical units gives a better representation of morbidity than comparisons between larger units such as cities or counties. Further analysis using specific rates will in all probability assist in identifying high risk groups in larger urban areas and will help in decision making regarding targeted health planning and interventions.

Table 2. Incidence rates of shigellosis by SES categories of census tracts for 7 cities of metropolitan Haifa

<table>
<thead>
<tr>
<th>SES category</th>
<th>Incidence rate (SD)</th>
<th>Kruskal-Wallis mean rank</th>
<th>Number of tracts in group</th>
<th>Kruskal-Wallis $\chi^2$</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>64.2 (40.1)</td>
<td>51.4</td>
<td>19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle</td>
<td>57.1 (40.4)</td>
<td>47.6</td>
<td>38</td>
<td>.0440</td>
<td>.803</td>
</tr>
<tr>
<td>High</td>
<td>70.0 (59.4)</td>
<td>51.6</td>
<td>42</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

SES, Socioeconomic status; SD, standard deviation.

References