Defining Core Gonorrhea Transmission Utilizing Spatial Data

Kyle T. Bernstein1,2, Frank C. Curriero3, Jacky M. Jennings4, Glen Olthoff2, Emily J. Erbelding2,5, and Jonathan Zenilman5

1 Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD.
2 Sexually Transmitted Disease Prevention Program, Baltimore City Health Department, Baltimore, MD.
3 Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD.
4 Department of Adolescent Medicine, School of Medicine, Johns Hopkins University, Baltimore, MD.
5 Department of Infectious Diseases, School of Medicine, Johns Hopkins University, Baltimore, MD.

Received for publication August 20, 2003; accepted for publication February 2, 2004.

Spatial distribution of repeat and singly occurring gonorrhea in Baltimore, Maryland, was examined to identify clusters of core transmitters. Gonorrhea reported between 2001 and 2002 was geocoded and mapped. Those with more than one gonorrhea infection separated by more than 14 days during the 2-year period were considered repeaters. Other cases were treated as isolated events. Six definitions of core transmission were examined by census tracts with the highest quintiles of 1) cases for 2001, 2) cases for 2002, 3) rates for 2001, 4) rates for 2002, 5) counts of repeaters over the 2-year period, and 6) proportion of total cases that were repeaters. Of the 6,108 gonorrhea cases analyzed, 9% were repeaters. Repeaters were more likely to be female and younger. Core areas identified by definitions based on overall disease burden agreed well with each other but had poor agreement with definitions based upon repeat infections. Repeaters clustered to a greater extent at smaller distances than did singly occurring gonorrhea cases. Repeat gonorrhea infections are prevalent in Baltimore and likely represent behavior consistent with core transmission. Census tracts of core transmission defined by geographic distribution of repeat infections may indicate foci of highest risk sexual behaviors and high transmission.

Keywords: disease transmission; gonorrhea; sexually transmitted diseases, bacterial; spatial distribution

Abbreviations: GIS, geographic information systems; STD, sexually transmitted disease.

The reproductive ratio of sexually transmitted diseases (STDs) in communities is a function of three components: the efficiency of transmission of disease between individuals, the rate of sexual partner exchange in a population, and duration of infectivity (1, 2). The sustainability of STD transmission in a population depends on core groups, defined as individuals who are consistently infectious and able to transmit disease to large numbers of sexual partners (3–5). For example, core groups may have a longer period of infectivity through repeat infections, thereby increasing the overall reproductive rate. Mathematical models of human immunodeficiency virus (6, 7) and gonorrhea (8, 9) epidemics have quantified the core group impact.

Although the core group is a valuable theoretical construct and mathematical tool (10–13), its implementation into public health practice has been hampered by a lack of standard operational definitions. Over 20 years ago, Rothenberg (14) analyzed reported gonorrhea in upstate New York and found distinct geographic areas of high morbidity, which he hypothesized reached levels sufficient to sustain epidemic transmission of disease. He defined core as geographic areas (census tracts) accounting for 50 percent of the morbidity in major urban areas (14). In Baltimore, Becker et al. (15) adapted computerized geographic information systems (GIS) to map reported gonorrhea cases, defining core as those census tracts in the highest rate quintile; they examined the spatial distribution of the burden and rate of gonorrhea infection, confirming Rothenberg’s findings. Jennings et al. (16) examined the spatial clustering in Baltimore City and found two distinct clusters of gonorrhea infections, which
they suggested may represent dense sexual networks with high disease prevalence. These examinations based defini-
tions of core groups on the local burden of disease (geospa-
tial core) and not on identification of individuals who may be
crucial to the continued transmission of disease (behavioral
core).

On a pragmatic level, previous analyses have shown that
individuals who become infected with gonorrhea multiple
times ("repeaters") account for a disproportionate share of
disease and would constitute a core group (17–22). A recent
analysis found that the incidence of repeat gonorrhea in the
two Baltimore City STD clinics from 1994 to 1998 was 4.28 per
100 person-years (17). Although the overall burden of
gonorrhea declined during this period, the proportion of
gonorrhea repeaters increased (17). Gonorrhea repeaters,
compared with those having isolated gonorrhea episodes,
have a longer duration of infectivity and may have a higher
rate of partner exchange. We believe that this population
may be an important focus of intervention and may be an
easily definable core group.

We compared the geospatial distribution of gonorrhea
repeaters and those with only one gonorrhea episode
reported to the Baltimore City Health Department between
2001 and 2002. We treat the gonorrhea repeaters as having characteristics consistent with a behavioral core group (a
core group based on behaviors such as multiple sexual part-
ners, increased duration of infectivity, etc.) and explore
whether these repeaters also exhibit geospatial core group
characteristics (spatial clustering). We developed and
compared several definitions of geospatial core and
compared the spatial clustering between the repeaters and
those with only one episode. We believe that this descriptive
analysis is a useful means by which local health practitioners
can gain a better understanding of the local epidemiology of
gonorrhea and better target interventions in an efficient and
effective manner.

MATERIALS AND METHODS

Data sources

Gonorrhea reporting by providers and laboratories is
mandated by law in Maryland. Diagnosis of gonorrhea is
based on the isolation of Neisseria gonorrhoeae from a
urethral, cervical, pharyngeal, or rectal swab or by positive
result from a swab or urine specimen tested with a licensed
nucleic acid amplification test (23). For reported episodes of
gonorrhea, the residential address, age, race/ethnicity,
gender, dates of diagnosis and reporting, and provider
reporting the case are captured in the STD Program’s elec-
tronic surveillance system. Data in this analysis represent
diagnosed gonorrhea infections from the two Baltimore City
Health Department STD clinics and all non-STD clinic
providers.

We evaluated all episodes of gonorrhea reported in 2001–
2002 to identify repeat infections. Episodes were matched
utilizing a unique patient identifier generated in the surveil-
lance system. Individuals who had multiple gonorrhea
episodes captured in the surveillance system between 2001
and 2002 that were more than 14 days apart were considered
repeaters. Those with only one gonorrhea episode or
multiple episodes that were separated by 14 days or less were
considered isolated cases. The treatment guidelines from the
Centers for Disease Control and Prevention recommend
single-dose (either oral administration or intramuscular
injection) therapy for gonorrhea infection that has greater
than a 98 percent efficacy (24). Given that gonorrhea cases
are treated with a single-dose regimen and that the incuba-
tion period is relatively short, we considered episodes
reported within 14 days or less of one another likely a duplic-
ate entry of the same episode.

Geocoding of gonorrhea cases

Residential addresses of repeat and isolated cases of
gonorrhea were geocoded with ArcGIS software (ESRI
(Environmental Systems Research Institute), Redlands,
California) using ESRI-supplied Maryland street basemaps.
Addresses that could not be geocoded were removed from
analysis. For gonorrhea repeaters with multiple addresses,
the residential address of the first gonorrhea infection was
utilized for analysis. The census tract for each case was
obtained by overlaying 2000 US Census boundary files with the
residential address. Aggregate census tract counts of the
number of isolated and repeat gonorrhea cases were gener-
ated. Census tract-specific rates were generated by dividing
the total reported gonorrhea cases per year by the 2000 US
Census population estimate.

Defining core

Six definitions of gonorrhea core groups were adapted from others (15) or created: 1) census tracts in the highest
quintile of reported gonorrhea cases in 2001, 2) census tracts in the highest quintile of gonorrhea rates in 2001, 3) census
tracts in the highest quintile of reported gonorrhea cases in
2002, 4) census tracts in the highest quintile of gonorrhea rates in 2002, 5) census tracts in the highest quintile of
gonorrhea repeater, and 6) census tracts with the highest
percentage of gonorrhea cases that were repeaters (the number of repeaters divided by the 2001–2002 average
number of gonorrhea cases).

Statistical analysis

Characteristics of gonorrhea repeaters and nonrepeaters were compared using $\chi^2$ test statistics. Areas defined as core
groups based on our six definitions were compared to assess
the level of agreement. For each of the six definitions, census
tracts were dichotomized into being in the highest quintile
compared with the four lower quintiles. Kappa values were
calculated between the different definitions (25). Addition-
ally, to examine agreement between the continuous values of
the six definitions of the core for each census tract, we gener-
ated correlation coefficients for comparison.

The extent of geographic clustering was established for the
repeaters and nonrepeaters using $K$-function analysis (26).
The estimated $K$ function, which measures the expected
number of further events within a range of distances $h$ from
observed events, is given by the following equation:


\[ K(h) = \frac{R}{n} \sum \sum I_h(d_{ij}), \]

where \( R \) is the area of the region of interest, \( n \) is the total number of events, \( d_{ij} \) is the distance between the \( i \)th and \( j \)th observed event, and \( I_h(d_{ij}) \) is the indicator function that is 1 if \( d_{ij} \leq h \) and 0 otherwise (26).

The \( K \) functions for the repeaters and nonrepeaters were compared to examine differences in the extent and resolution of geographic clustering, testing the null hypothesis that \( K_{\text{repeaters}}(h) = K_{\text{nonrepeaters}}(h) \). Monte Carlo simulations were used to generate confidence envelopes for the difference in \( K \) functions, \( K_{\text{repeaters}}(h) - K_{\text{nonrepeaters}}(h) \) for the range of distances \( h \), based on randomly permuting repeat and nonrepeat location labels to provide the corresponding distribution under the null hypothesis (27). The fifth and 95th percentiles for the difference in \( K \) functions at each distance \( h \) represent 95 percent confidence envelopes in which to assess significant differences in geographic clustering between the repeat and non-repeat gonorrhea infections. The R-language statistical computing environment (28) with the SPLANCS contributed software package (http://www.r-project.org) was used for the \( K \)-function analysis. SAS version 8.2 software (SAS Institute, Inc., Cary, North Carolina) and STATA version 8 software (StataCorp LP, College Station, Texas) were utilized to calculate kappas and correlation coefficients.

RESULTS

Between January 1, 2001, and December 31, 2002, 9,910 gonorrhea episodes were reported to the STD Surveillance Unit of the Baltimore City Health Department. Of these, 1,512 episodes (15.3 percent) represent repeat infections. The number of unique individuals infected with gonorrhea during this 2-year period was 9,097, of which 699 (7.7 percent) were gonorrhea repeaters. Among the 699 gonorrhea repeaters, 606 (86.6 percent) had two episodes of gonorrhea during the 2-year study period, 76 (10.9 percent) had three, 13 (1.9 percent) had four, and four (0.5 percent) had five gonorrhea episodes.

Geocoding results

Of the 9,097 individuals with a gonorrhea infection (repeaters and nonrepeaters), 6,595 (72.5 percent) had a residential address listed. A total of 6,418 (97.3 percent) gonorrhea cases with residential addresses were geocoded; 175 (2.7 percent) could not be matched to a valid geocodable address. Among those 6,418 cases that could be geocoded, 485 (5.3 percent) resided outside the geographic boundaries of Baltimore City and were removed from further analysis. Subjects included in this analysis were therefore 6,108 gonorrhea cases (individuals) who resided in Baltimore City and had a valid address, of which 550 (9.0 percent) were repeaters.

Individuals who could not be geocoded differed significantly from those who could. Gonorrhea cases with a geocodable address were more likely to be older (\( p < 0.05 \)), male (\( p < 0.05 \)), and African American (\( p < 0.05 \)). Similar differences were found when stratified by repeater status or by year.

The characteristics of the 599 repeaters and 5,558 non-repeaters are shown in table 1. Gonorrhea repeaters with valid Baltimore City residential addresses were more likely to be younger (mean age: 21.7 vs. 24.7 years) and more likely to be female. Repeaters were more likely to have been reported to the Baltimore City Health Department by community-based organizations and school-based clinics, and they were less likely to be reported by private medical providers than were nonrepeaters. Among the repeaters, the median time to first reinfection was 134 days (range: 20–609 days).

Maps operationalizing our definitions of core groups are shown in figures 1, 2, and 3, and the kappa statistics and correlation coefficients from the comparisons of the six dichotomized definitions of core groups are shown in table 2. The agreement analysis suggests that the core group definitions based on the 2001 rate and count and on the 2002 rate and count data agree well. Since the definitions based on gonorrhea counts and rates for 2001 and 2002 showed adequate agreement, we show only the 2002 rate map (figure
Additionally, the definitions of core groups based on the count and percentage of repeaters per census tract agreed well with each other (figures 2 and 3). However, the definitions of core groups based on simple disease burden (2001 and 2002 count and rate) did not agree or correlate well with the definitions based on repeat infections.

The S shape of the $K$ functions (values not shown) for both the gonorrhea repeaters and nonrepeaters suggests classic...
spatial clustering (27). This phenomenon was further examined, and the differences between the $K$ function for the repeaters and nonrepeaters are shown in figure 4. The simulations at 100, 200, 300, and 400 iterations showed similar results, and only the 100-iteration run is displayed. At distances where the difference in $K$ functions exceeds the 95 percent envelopes, differential spatial clustering is found. At small distances (approximately 0.33 mile or 0.53 km), there appeared to be a significantly higher level of spatial clustering of the gonorrhea repeaters compared with the nonrepeaters. This difference was not significant at larger distances.

**DISCUSSION**

Using surveillance data, we found that 9 percent of the individuals with gonorrhea and a valid Baltimore City address had multiple gonorrhea episodes between 2001 and 2002. Utilizing repeat gonorrhea infections as a marker of high-risk sexual behavior, we found that the geographic area of high overall gonorrhea burden did not necessarily correspond to areas with a high burden of repeat infections. Furthermore, at small distances, the gonorrhea repeaters seem to exhibit a higher level of spatial clustering than those with isolated gonorrhea events (nonrepeaters).

GIS technology has been shown to be a valuable tool in understanding STD transmission and in the design of intervention (29, 30). Mapping incident bacterial STDs can provide clues to epidemiologic disease patterns and provide insight into the possible risk factors for disease acquisition beyond the limited data that are often collected at the local level. This is facilitated by increased local health department GIS capacity.

<table>
<thead>
<tr>
<th>TABLE 2. Kappa statistics* (correlation coefficients†) for comparisons of six definitions of geospatial core of gonorrhea in 200 census tracts, Baltimore, Maryland, 2001–2002</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>2001 count</td>
</tr>
<tr>
<td>2001 rate</td>
</tr>
<tr>
<td>2002 count</td>
</tr>
<tr>
<td>2002 rate</td>
</tr>
<tr>
<td>Repeat count</td>
</tr>
<tr>
<td>% repeat</td>
</tr>
</tbody>
</table>

* Kappas were calculated for dichotomized definitions (highest quintile vs. not in highest quintile).
† Correlation coefficients were calculated as continuous variables.
While the STD literature has described the values of the core group from a conceptual, disease-control framework, there has been little consensus on how these core groups can be defined (3, 4, 14, 15, 18, 31). Early work has shown the spatial patterning of reported gonorrhea cases, suggesting highly focal hyperendemicity in certain areas (11, 15). However, since these conceptualizations of core groups did not include any incorporation of a behavioral component, they may not directly map to risk behaviors that may facilitate increased disease transmission, such as multiple sexual partners, lack of treatment for gonorrhea infection, and concurrent sexual partners.

We believe that geographically defined areas with a high incidence of repeat infections may be a better approximation of the core groups, as they combine geographic and behavioral components that can be obtained from routinely collected surveillance data. From a practical standpoint, the methods utilized are relatively simple to implement and have promise in the allocation of resources for intervention and prevention activities at the local level.

In Baltimore City, the mapping of disease patterns has been effective in syphilis elimination efforts, by focusing outreach and screening activities on geographic areas implicated in disease transmission (30). We believe these successes can be replicated for gonorrhea control. However, screening and interventions based on areas of high disease burden may have a lesser effect than focused attention to core areas defined through recidivism. The utility of intervention at the level of gonorrhea repeaters is largely unknown and may represent a productive disease control direction.

At short distances, the gonorrhea repeaters appeared to cluster more than the nonrepeaters. At larger distances, the extent of clustering between the repeaters and nonrepeaters did not differ. Therefore, repeaters tended to reside in closer proximity to other repeaters, compared with individuals having isolated gonorrhea events. The tighter clustering of repeaters at small distances supports our hypothesis. These geographic pockets of individuals exhibiting behavioral characteristics consistent with core group transmission may be valuable foci of intervention.

We found that females were more likely to be gonorrhea repeaters than males. This finding is novel considering that females are more likely to have asymptomatic gonorrhea infections compared with males (32, 33) and are in turn less likely to have a gonorrhea infection diagnosed and reported to the local health department. It is possible that females with repeat gonorrhea infections are more likely to engage in exchanges of sex for drugs or money. Mehta et al. (17) found that, among Baltimore City STD clinic clients, females were more likely to have a repeat gonorrhea infection; however, after multivariate adjustment, this finding was not significant. The data utilized in this analysis did not include behavioral information about the cases, such as engagement in exchange sex. Additionally, since it is the policy of the Baltimore City Health Department to ensure treatment for all reported female gonorrhea cases but not male cases, it is possible that male partners of treated females never receive treatment themselves and in turn reinfect their partner.

There are several limitations to our analysis. First, over a quarter of the reported gonorrhea cases during our 2-year analysis period could not be geocoded. These cases that could not be mapped were older and more likely to be male. This may result from the Baltimore City Health Department policy to confirm treatment for all reported female gonorrhea cases but not male gonorrhea cases. Since active field efforts are used to confirm gonorrhea treatment, female cases (who are more likely to be younger than males) are more likely to have a complete and useful address on record because of program activities. While the sample used in this analysis differed from those without addresses based on demographics, we have no reason to believe that the repeaters and nonrepeaters not geocoded differed spatially from those geocoded.

A recent population-based study of chlamydia and gonorrhea infections found that 5.3 percent of Baltimore residents aged 18–35 years had an untreated gonorrhea infection (34). Furthermore, based on this analysis, it was estimated that the true burden of gonorrhea infections in Baltimore far exceeded the number reported to the Baltimore City Health Department (34). The study presented here was limited to reported gonorrhea infections and likely underestimated the true burden of repeat infections in the community. Both low-grade cervical infections in women and asymptomatic carriage of gonococci in the pharynx are not likely to result in receipt of a diagnostic test and subsequent reporting to the Health Department. Thus, these data are limited in their ability to provide a population-level estimation of the burden.
of disease. However, much of local public health policy is based on the burden of reported diseases. The objective of this report was to apply a novel approach to the analysis and interpretation of locally available data.

This analysis was limited to the residential address of reported gonorrhea cases in Baltimore City, Maryland, and not to the location of sexual partner recruitment. However, Zenilman et al. (31) have shown that, in Baltimore City, individuals with gonorrhea tended to live relatively close to their sexual partners, with females a median of 547 m from their partners and males a median of 339 m from their partners. This finding suggests that, in Baltimore City, gonorrhea cases select sexual partners that reside close to their residence, and the location of sexual partner recruitment is focused near that residence. The analysis presented here utilized data that are easily accessible at the local level and, while the location of sexual partner recruitment may be a more useful measure for disease interruption, we believe residential address is an appropriate proxy measure for use in public health intervention.

Residential addresses of gonorrhea cases were not consistently entered into the Baltimore City surveillance system prior to 2001. As a result, this analysis was restricted to gonorrhea reported after 2001. Substantially more repeaters would have likely been identified with a larger study window, but to increase the proportion of cases that could be geocoded, we had to restrict the analysis to 2001–2002. Third, due to the limited data collected in gonorrhea reporting, analyses comparing the repeaters and non-repeaters were restricted to demographic characteristics.

As the local capacity to utilize GIS technologies improves, the complexity of the application of spatial data to drive public health interventions will be a challenge. Here, we have shown that areas that may be targeted for gonorrhea intervention differed on the basis of definitions of core groups drawn from total disease burden compared with repeat infections. Focused efforts to reduce gonorrhea morbidity that target core transmitters may have a larger impact, particularly in areas heavily burdened by gonorrhea such as Baltimore City, than attempting to intervene with all reported cases. Additionally, the provision of partner services and notification may be more cost-effective if applied to populations of repeaters. Evaluations of the spatial distribution of STDs can guide both prevention and intervention efforts, as well as provide additional clues to the local transmission of disease.

ACKNOWLEDGMENTS

This work was supported by National Institute of Allergy and Infectious Disease grants T32AI50056 (K. T. B.) and K24AI01633 (J. Z.).

REFERENCES