Published in final edited form as:

Pediatr Infect Dis J. 2008 November ; 27(11): 981–985. doi:10.1097/INF.0b013e31817acf4f.

Spatial and Temporal Clustering of Kawasaki Syndrome Cases

Annie S. Kao, PhD^{*}, Arthur Getis, PhD[†], Stephanie Brodine, MD^{*}, and Jane C. Burns, MD[‡]

^{*} Division of Epidemiology, Graduate School of Public Health, San Diego State University, San Diego

[†] Department of Geography, San Diego State University, San Diego

[‡] Department of Pediatrics, Rady Children's Hospital San Diego and UCSD School of Medicine, La Jolla, CA

Abstract

Background—The etiology of Kawasaki syndrome (KS) remains unknown despite 30 years of intensive search for an agent. Epidemiologic clues to a possible infectious etiology include the seasonal distribution of cases, the previous occurrence of epidemics, the clinical features of the syndrome that mimic other infectious rash/fever illnesses in children, the self-limited nature of the illness, and the peak age incidence in the toddler years.

Methods—We examined the epidemiology and spatial and temporal distribution of KS cases in San Diego County, California during the 6-year period from 1998 to 2003. Clustering in space and time was analyzed using geo-referenced data with the K-function, the local G-statistic, and Knox statistic.

Results—A total of 318 patients were identified through active surveillance. The overall annual incidence was 21.7/100,000 in children <5 years, with rates in whites, white Hispanics, and Asian/Pacific Islanders of 15.3, 20.2, and 45.9/100,000, respectively. The Knox test showed significant clustering of cases within the space-time interval of 3 km and 3–5 days.

Conclusions—This is the first study of KS cases to use geo-referenced point pattern analysis to detect spatial and temporal clustering of KS cases. These data suggest that an infectious agent triggers the immunologic cascade of KS.

Keywords

epidemiology; coronary artery aneurysms; vasculitis

The etiology of Kawasaki syndrome (KS), the most common cause of acquired heart disease in children, has eluded investigators for more than 30 years.¹ Despite its clinical similarity to other infectious diseases, exhaustive search for the causative agent has been thus far unsuccessful. This acute, self-limited vasculitis results in permanent coronary artery damage in up to 25% of untreated children. High dose intravenous gamma globulin reduces the risk of coronary artery aneurysm to 3–5% if administered early in the course of disease.² However, without a specific diagnostic test, affected children may be difficult to recognize, and delayed diagnosis and treatment continue to result in potentially preventable morbidity and mortality.³ Thus, there is urgency to the quest for the etiologic agent. Historically, study of the distribution of human diseases within a community has led to valuable insights into their

Address for correspondence: Jane C. Burns, MD, Department of Pediatrics, UCSD School of Medicine, 9500 Gilman Dr, La Jolla, CA 92090-0830. jcburns@ucsd.edu.

A. S. Kao is currently at Community Epidemiology Branch, San Diego County Health and Human Services Agency, San Diego, CA.

Presented in part at the 8th International Kawasaki Disease Symposium, San Diego, CA, February 2005.

causation. To further test the hypothesis that the immunologic reaction that we recognize as acute KS is triggered by an infectious agent, we analyzed KS cases in a circumscribed region with respect to time and place. Cluster analysis of this geo-referenced data suggested a nonrandom distribution of cases, lending further weight to the hypothesis that acute KS is triggered by an infectious agent.

METHODS

Case Ascertainment

Title 17 of the California Code of Regulations (CCR), requires health care providers in California to report KS cases to their local health departments. The San Diego County Health and Human Services Agency, Community Epidemiology Branch collects data on reported patients with a diagnosis consistent with KS. Medical records are reviewed quarterly, and cases are verified at the end of each calendar year. At 1 site, where the majority of KS cases are hospitalized, active surveillance is maintained throughout the year. Continual case finding and medical records reviews ensure optimal ascertainment of diagnosed cases.

During the period of this study, 4 hospitals in San Diego County (a pediatric health center, a comprehensive military hospital, a full-service hospital and trauma center, and a health maintenance organization) admitted KS patients, with greater than 90% of the KS cases in the county admitted to the pediatric health center. All patients hospitalized between January 1, 1998 and December 31, 2003 at any of these 4 hospitals were identified by ICD-9 code 446.1 (KS or mucocutaneous lymph node syndrome) in the discharge diagnosis. Each patient was counted only once, even if there were multiple hospitalizations for that individual. Medical records were reviewed by the health department and by the Kawasaki Disease Research Center at UCSD/Rady Children's Hospital San Diego. Variables used for descriptive epidemiology were abstracted from the hospital records. Geo-spatial data included addresses, zip codes, and latitude and longitude data collected from the GIS mapping software.

Patients who met the American Heart Association (AHA) case definition for KS and whose families resided in San Diego County were included.¹ Patients with at least 5 days of fever and fewer than 4 clinical criteria with coronary artery abnormalities were classified as having incomplete KS.

Descriptive Data Analysis

We examined the demographic characteristics of KS patients admitted during the 6-year period from January 1998, through December 2003. Population data for children <5 years of age were obtained from the San Diego Association of Governments (SANDAG),⁴ based on estimates from the 2000 U.S. Census. Incidence of KS was calculated per 100,000 children <5 years of age for each year, by age, gender, and race/ethnicity and frequencies were analyzed using EPI Info 6.04c (CDC, Atlanta, GA).⁵ Continuous variables were compared using Student *t* tests. Categorical variables and proportions were compared using the χ^2 test or the Fisher exact test (2-tailed). Statistical analyses were performed using EPI Info 6.04c (CDC, Atlanta, GA) and SPSS Version 10.0.⁶

Spatial and Temporal Analysis

The residence of each KS case was identified and sorted by zip code area to evaluate spatial distribution. Point pattern analysis of geo-referenced data was used for identifying statistically significant clusters of KS cases. The point pattern techniques used in this study included the K-function,⁷ local G-statistic [$G_i^*(d)$],⁸ and Knox statistics for space-time clustering.⁹ All assumptions for these tests were reviewed before reporting the statistical results.

Latitude and longitude data for the primary residence of each case and for the centroid of each zip code area were converted to Universal Transverse Mercator (UTMs) coordinates, and the distance (in km) between cases or between zip codes was compared. Incidence rates for each zip code were calculated using population data provided by the 2000 U.S. Census Bureau. Besag-Newell corrections were used to normalize the crude rates.¹⁰ Maps illustrating disease frequency (counts) and crude rates were created using Arcmap 8.2 (Environmental Systems Research Institute).

K-Function

The K-function analysis considers all combinations of pairs of KS cases and compares the distance $[L(d)]$ between the primary residence of the observed pairs with a simulated data set that was generated by examining 99 permutations of pairs of randomly generated points over the study area. The K-function takes into account the number of cases and the area of the region of interest (San Diego County). The spatial pattern of pairs of KS cases was compared with the pattern of control pairs derived from the general population. If the observed distance between KS pairs fell above or below the simulated distance between pairs, within a 95% envelope around the simulated pattern, the null hypothesis of complete spatial randomness was rejected. An observed $L(d)$ above the maximum simulated $L(d)$ indicates clustering of cases at that distance. The maximum study distance was set at 15 km, with 30 increments so that the observed $L(d)$ and confidence interval were calculated for every 0.5 km.

Local $G_i^*(d)$ Statistic

This statistic considers the location of the varying rates of the disease and identifies the clustering of these rates at levels greater than expected and detects spatial clustering that may not be evident using global statistics (K-functions). Generally, this statistic can be used to detect local spatial clustering around individual locations that would not be detectable by more general global statistics. The $G_i^*(d)$ requires that the observed value is included in the calculation for spatial clustering and the statistical tests are based on the normal distribution. The local G_i^* statistic was calculated for a specified distance of 5 km from each KS case at 1-km increments for zip code areas with at least 2 KS cases.⁸

Knox Statistic

The Knox statistic was used to test whether there was clustering within a defined distance and time period. Pairs of KS cases within a specified space and time interval were counted and compared with the expected number of pairs within the same interval. A P value <0.05 , based on the Poisson distribution, indicated significant space-time clustering within the given time and distance intervals. Because the incubation period for many infectious diseases does not extend beyond 10 days and because of our limited sample size, we examined space-time clustering at 1, 3, 5, 7, and 10 days after disease onset for each KS pair.

RESULTS

KS Incidence

During the study period, there were 318 KS cases: 274 cases (86%) met the classic case definition with fever plus 4/5 criteria and 44 cases (14%) had incomplete KS. The overall annual incidence for KS was 21.7 cases per 100,000 children <5 years. The incidence increased from 17.3 per 100,000 in children <5 years in 1998 to 22.8 in 2003, although this difference was not statistically significant.

Demographic characteristics of children <5 years hospitalized with KS between January 1, 1998 and December 31, 2003, representing approximately 80% of the cases are shown in Table 1.

Race and Ethnicity

Overall, Asian and Pacific Islander (API) children and children of mixed races <5 years of age had the highest KS rate (Table 1, Fig. 1). These rates were significantly higher compared with white Hispanics, African American, and white children. Non-Hispanic white children had the lowest rates. API children showed consistently higher incidence rates from 1998 to 2002 compared with the other race/ethnicity groups (Fig. 1). The reason for the decreased incidence in this group in 2003 is unclear. The rates for African American children were based on very small numbers in the numerator for each year (3–5) because of the small African American population in San Diego County.

Geographic Variation

We calculated the rates for KS cases in 24 of 104 zip code areas that had 5 or more KS cases <5 years of age (Fig. 2). The rates of KS cases varied by region. The 2 zip code areas in the central region with the highest rates of KS cases also had a large API population (50% and 57%). However, the other regions of high incidence could not be explained by the distribution of the API population. Cases residing in the 2 highest incidence zip codes in the north coastal area were predominantly white with only 15% API, and cases in the 3 highest zip codes in the southern region were mostly white children of Hispanic origin with only 30.8% API. Thus, the uneven distribution of rates throughout the county could not be explained by either the distribution of children <5 years or the distribution of the API population, a group which has the highest incidence of KS.

K-Function

There were a total of 311 KS cases in this analysis because addresses for 7 cases were unavailable. At a distance >2 km (dark vertical line, Fig. 3), the observed distance fell within the confidence interval for the expected distance. For distances <2 km, the observed distance was above the confidence interval, which indicated clustering of cases within 2 km (Fig. 3). Clustering was most marked between 1–1.1 km (arrow, Fig. 3) at which point the distance between the observed and the upper bound of the confidence interval was greatest ($P < 0.01$).

Local G_i^* Statistic

This analysis looks at clustering of cases within a specified distance within the county. Two areas encompassing multiple zip codes showed significant clustering of cases within 1 km of each other (Fig. 4). To examine whether or not the apparent clustering was simply because of a skewed distribution of the Asian population in San Diego County, we looked at the percent API of the 4 zip codes with the highest clustering of KS cases versus other regions of the county. These 4 zip codes ranked 1st, 19th, 24th, and 66th of 104 zip codes in the county with respect to numbers of API residents.⁴ Thus, higher numbers of API might explain the clustering in only 1 of the 4 zip code areas.

Knox Statistic for Space-Time Clustering

The Knox statistic was used to test for significant clustering within a defined distance and time period. The distance in km between each case and the next case in time was compared. Given the results of the K-function and local G_i^* statistic analyses, we chose to analyze a space-time window of 1 and 3 km and intervals between cases of 1, 3, 5, 7, and 10 days. No significant space-time clustering was detected for intervals greater than 5 days (data not shown). For a distance of 1 km, which represents a neighborhood area of 6–7 blocks, there were very few

cases. For a distance of 3 km and a time interval of 3 or 5 days, 9 and 12 pairs of points, respectively, fell within the specified space-time window (expected 4.3 and 6.9 pairs of points, respectively; Table 2). Thus, there was significant clustering of cases within the space and time interval of 3 km and 3–5 days.

DISCUSSION

This is the first study to use geo-referenced data in a space-time cluster analysis. The Knox test showed significant clustering of cases within the space and time interval of 3 km and 3–5 days. Statistically significant spatial and temporal clustering occurred in 2 areas of the county within the space and time interval of 3 km and 3–5 days.

One limitation is that our study does not exclude the possibility of space-time clustering over a lesser distance. However, because of the small sample size observed for distances <3 km, our analysis was limited to categories with sufficient numbers. In addition, both space-time clustering analysis performed at 3 km and 3–5 days showed significant numbers in the observed compared with the expected numbers making this finding less likely a result of chance. Another limitation of the study was the inability to take into account population density in the K-function analysis, which could certainly be a factor contributing to clustering of cases.

San Diego County is the third most populated county in California and includes the eighth largest metropolitan area in the United States. The incidence reported in this study is slightly higher than the U.S. estimate (17.1/100,000), reported by the Centers for Disease Control and Prevention.¹¹ However, it is substantially lower than the incidence of KS reported in Hawaii (47.7/100,000 children <5 years),¹² Taiwan (54.9/100,000 children <5 years),¹³ and Japan (137.7/100,000 children <5 years).¹⁴ Additionally, unlike the increasing incidence in Japan, the KS incidence in San Diego County did not change significantly during the study period.

The county is racially and ethnically diverse; it has an estimated population close to 3 million residents characterized as 55.0% non-Hispanic whites, 26.7% Hispanic whites, 5.5% non-Hispanic blacks, and 12.7% non-Hispanic API and other (data from 2000 Census). Although KS has been reported in most ethnic groups, the disease is overrepresented among Asian Pacific Islander children and children of mixed races. Although the API population <5 years old represents approximately 14% of the population in San Diego County,⁴ they accounted for 27.9% of the cases of KS in this study.

This is the first comprehensive analysis of KS epidemiology using surveillance data in a community with a large Hispanic population. Since 1990, the Hispanic population has increased in San Diego County by 47%.⁴ Although past studies have reported Hispanic children to have a lower incidence of KS (6.5–9.6/100,000 children <5 years),¹⁵ this study found a higher incidence than previously reported (20.21/100,000 children <5 years). Research by our group has shown Hispanics to have differences in health-seeking behavior, higher barriers to accessing health services, and delayed diagnosis as a risk factor for developing coronary artery abnormalities.¹⁶ Moreover, a recent study using data from a passive surveillance system found that there was a significant association between Hispanic ethnicity and development of CAA resulting from delayed treatment.¹⁷

The KS rates varied regionally across the county and across zip code areas. This study confirmed previous reports that APIs have a higher frequency of KS and that areas with a larger Asian population have higher incidence rates. The zip code areas with the highest KS incidence also had the largest proportion of API (≥50%) compared with other zip codes across the county. Further evaluation of the spatial distribution using point pattern analysis for each KS case identified significant clustering of KS cases within a critical distance of 1 km. The spatial clustering in the present study may suggest an environmental trigger such as an infectious agent

that results in clinical KS only in genetically susceptible individuals. Seasonality of KS with winter/spring and summer peaks has been previously described.¹⁸ This new observation of clustering of cases on a smaller time scale lends further weight to the infectious trigger hypothesis for KS.

Associations have been observed between antecedent respiratory illness and KS,¹⁹ and researchers have proposed that an agent causing KS could first infect the upper respiratory tract before triggering a systemic immunologic response.²⁰ Increased IgA plasma cell infiltration of the upper respiratory tract and coronary arteries in KS patients further support the respiratory route as a potential portal of entry for the causative agent.²¹ The presence of IgA-secreting plasma cells in the upper respiratory tract mimics the response seen in autopsies of children who died of known viral respiratory infection such as respiratory syncytial virus.

The new finding of KS clusters limited in space and time supports the hypothesis of an infectious etiology and suggests that focused study of KS clusters may yet yield clues to the inciting agent.

Acknowledgments

Supported in part by grant NIH-NHLBI K24 HL074864 (to J.C.B.).

References

1. Newburger JW, Takahashi M, Gerber MA, et al. Diagnosis, treatment, and long-term management of Kawasaki disease: a statement for health professionals from the Committee on Rheumatic Fever, Endocarditis and Kawasaki Disease, Council on Cardiovascular Disease in the Young, American Heart Association. *Circulation* 2004;110:2747–2771. [PubMed: 15505111]
2. Newburger JW, Takahashi M, Beiser AS, et al. A single intravenous infusion of gamma globulin as compared with four infusions in the treatment of acute Kawasaki syndrome. *N Engl J Med* 1991;324:1633–1639. [PubMed: 1709446]
3. Rowley AH. Finding the cause of Kawasaki disease: a pediatric infectious diseases research priority. *J Infect Dis* 2006;194:1635–1637. [PubMed: 17109334]
4. San Diego Association of Governments (SANDAG), constructed from U.S. Census Bureau's 2000 Census Summary Files 1 and 35.
5. Epi Info. [computer program]. 6.04c ed. Atlanta, GA: CDC; 1995.
6. SPSS. [computer program]. 10. Chicago, IL: 1999.
7. Ripley, BD. *Spatial Statistics*. New York, NY: Wiley; 1981.
8. Getis A. The analysis of spatial association by use of distance statistics. *Geogr Anal* 1992;24:189–206.
9. Knox G. The detection of space-time interactions. *Appl Stat* 1964;13:25–29.
10. Besag J. The detection of clusters in rare diseases. *J R Stat Soc A* 1991;154:143–155.
11. Holman RC, Curns AT, Belay ED, Steiner CA, Schonberger LB. Kawasaki syndrome hospitalizations in the United States, 1997 and 2000. *Pediatrics* 2003;112(3 Pt 1):495–501. [PubMed: 12949272]
12. Holman RC, Curns AT, Belay ED, et al. Kawasaki syndrome in Hawaii. *Pediatr Infect Dis J* 2005;24:429–433. [PubMed: 15876942]
13. Lue HC, Philip S, Chen MR, Wang JK, Wu MH. Surveillance of Kawasaki disease in Taiwan and review of the literature. *Acta Paediatr Taiwan* 2004;45:8–14. [PubMed: 15264699]
14. Yanagawa H, Nakamura Y, Yashiro M, et al. Incidence survey of Kawasaki disease in Japan: the nationwide surveys of 1999–2002. *Pediatr Int* 2006;48:356–361. [PubMed: 16911079]
15. Bronstein DE, Dille AN, Austin JP, Williams CM, Palinkas LA, Burns JC. Relationship of climate, ethnicity and socioeconomic status to Kawasaki disease in San Diego County, 1994 through 1998. *Pediatr Infect Dis J* 2000;19:1087–1091. [PubMed: 11099092]

16. Wilder MS, Palinkas LA, Kao AS, Bastian JF, Turner CL, Burns JC. Delayed diagnosis by physicians contributes to the development of coronary artery aneurysms in children with Kawasaki syndrome. *Pediatr Infect Dis J* 2007;26:256–260. [PubMed: 17484225]
17. Belay ED, Maddox RA, Holman RC, Curns AT, Ballah K, Schonberger LB. Kawasaki syndrome and risk factors for coronary artery abnormalities: United States, 1994–2003. *Pediatr Infect Dis J* 2006;25:245–249. [PubMed: 16511388]
18. Burns JC, Cayan DR, Tong G, et al. Seasonality and temporal clustering of Kawasaki syndrome. *Epidemiology* 2005;16:220–255.
19. Bell DM, Brink EW, Nitzkin JL, et al. Kawasaki syndrome: description of two outbreaks in the United States. *N Engl J Med* 1981;304:1568–1575. [PubMed: 7231502]
20. Rowley AH, Baker SC, Shulman ST, et al. Detection of antigen in bronchial epithelium and macrophages in acute Kawasaki disease by use of synthetic antibody. *J Infect Dis* 2004;190:856–865. [PubMed: 15272416]
21. Rowley AH, Shulman ST, Mask CA, et al. IgA plasma cell infiltration of proximal respiratory tract, pancreas, kidney, and coronary artery in acute Kawasaki disease. *J Infect Dis* 2000;182:1183–1191. [PubMed: 10979916]

FIGURE 1.

Kawasaki syndrome rates by race/ethnicity in children <5 years, San Diego, California, 1998–2003. Rates were based on estimates derived from the SANDAG and Census 2000. Asian rates include Asian Pacific Islander (API)/Other/Mixed to match the denominator data provided by SANDAG.

FIGURE 2.
Distribution of Kawasaki syndrome cases by zip code and age-specific rates, San Diego,
California, 1998–2003.

FIGURE 3.

Graph of K-function analysis. Arrow indicates the point at which the distance between observed cases (dark line) is maximally divergent from the upper bound of the confidence interval (gray line) for the expected distance between cases.

FIGURE 4. Map of $G_i^*(d)$ results showing 2 areas with significant spatial clustering of cases within 1 km of each other.

TABLE 1

Demographic Characteristics of Children <5 Years of Age Residing in San Diego County Diagnosed With Kawasaki Syndrome During the 6-Year Period From January 1, 1998 to December 31, 2003

Characteristic	No. (%)	Rate per 100,000 Children <5 Yrs
Age category		
All children <5 yrs of age	258 (100%)	21.72
<6 mo	26 (10.1)	NA *
6–12 mo	39 (15.1)	NA
1–4 yr	193 (74.8)	NA
Sex		
Male	157 (60.9)	26.03
Female	101 (39.1)	17.27
Race/ethnicity		
White (non-Hispanic)	75 (29.1)	15.32
Black (non-Hispanic)	19 (7.4)	25.10
Hispanic	92 (35.7)	20.21
Asian/Pacific Islander/other	72 (27.9)	45.91

* Denominator not available.

TABLE 2

Results of the Knox Test Showing Number of Cases Within the Space-Time Intervals of 1 and 3 km and 3–5 Days

Distance	Time	Observed	Expected Number	<i>P</i>
1 km	3 d	0	0.79	0.55
1 km	5 d	2	1.18	0.33
3 km	3 d	9	4.33	0.03
3 km	5 d	12	6.90	0.04