The Spatial Distribution of Ross River Virus Infections in Brisbane: Significance of Residential Location and Relationships with Vegetation Types

Andreas MUHAR*1*2, Patricia E.R. DALE*1, Lukman THALIB*3 and Eri ARITO*1

*1 Faculty of Environmental Sciences, Griffith University, Brisbane, Australia.
 *2 Institut fur Freiraumgestaltung und Landschaftspflege, Universitat fur Bodenkultur, Vienna, Austria.
 *3 School of Public Health, Griffith University, Brisbane, Australia.

Abstract

For the study area of Brisbane City (population 800,000), Australia, 2160 cases of Ross River virus (RRv) infections from the years 1991 to 1996 were geocoded. Their spatial distribution was investigated using census data at the suburb level (162 units). Infection rates have been calculated and adjusted to the age distribution within each suburb. Signed chisquare tests showed that a large number of suburbs has significantly high or low infection rates. Using Principal Component Factor analysis and regression, a relationship was shown between the proportion of wetlands and bushland in a suburb and the infection rate of RRv. Although flight ranges of up to 50 km have been reported for the major vector species *Aedes vigilax* (Skuse), this study indicated that RRv infection risk is significantly high relatively close to mosquito habitats. There were significant differences in the infection rate of RRv between years, however the spatial associations did not appear to differ.

Key words: Ross river virus, spatial distribution, vegetation, mosquitoes, arbovirus

Introduction

The risk of contracting mosquito-borne disease is a function of the spatial and temporal pattern of vector breeding habitats and also of the interactions between vector and people. Relating ecological and entomological patterns is a useful tool for predicting disease risk and is a potentially important management tool ¹⁰. Geographic Information Systems (GIS) have emerged in recent literature as innovative tools to manage disease and particularly vector-borne diseases^{2, 3, 4, 5)}. Identifying the factors which specifically influence the risk of disease using GIS can assist management and is useful in assigning treatment priorities^{6, 7)}. Other developments include modelling the dynamics of disease spread (for example, Bogs et al.)⁸⁰. In this paper we present the results of analysing the distribution of cases of an arbovirus disease (Ross River virus) and its relationship with vegetation.

Ross River virus (RRv) is a mosquito-borne alphavirus. It is a polyarthritic disease with no specific treatment. It is evidenced by a range of symptoms including viraemia, aches in the joints, lassitude, rash and headaches. Some or all of these symptoms may be present. No known cases of death have occurred from the disease.

Received Jun. 4 1999/Accepted Oct. 7 1999 Reprint requests to: Dr P DALE Australian School of Environmental Studies Faculty of Environmental Sciences Griffith University Nathan Queensland 4111 Australia. TEL: (61-7)3875-7136 FAX: (61-7)3875-7459 Email: p.dale@maibox.gu.edu.au Although it is not the only arbovirus disease in Australia, RRv is currently the most prevalent one, with up to 8000 cases reported annually⁹. Because of the variable and sometimes minor symptoms of this disease the real number of infections is likely to be much higher than the reported figures, even though the diagnostic experience of general practitioners has increased significantly in the past decade.

In Queensland, two mosquito species are seen as the main vectors of RRv, the saltmarsh mosquito *Aedes vigilax* (Skuse) and the freshwater species *Culex annulirostris* Skuse. Other species have been identified as possible vectors under laboratory conditions. RRv is not specific to humans, in fact a main pool for the virus seems to be the macropods (kangaroos and wallabies), and RRv antibodies have been found in many other mammal species including cattle and rats. Transmission from human to human may be possible, but is thought to be likely only during epidemics¹⁰.

The concept of landscape epidemiology, as developed by Pavlovsky seems to be a promising approach to assess the spatial distribution of infection risk of a disease such as RRv with its close links between vectors and natural features¹¹). Landscape epidemiology studies the relationships between the structure and zoonoses of a landscape and the occurrence of diseases both under natural conditions and in the cultural landscape, with the goal of developing management strategies based on this knowledge. A first conceptual approach to mosquito risk modelling using the principles of landscape epidemiology was developed by Barnes et al.¹²). However, until recently, little research work has been conducted in Queensland to identify local differences in the spatial distribution of the RRv infection risk with regard to natural features. Mottram et al. investigated the prevalence of mosquito species as well as the distribution of RRv cases in selected Brisbane suburbs, Queensland, at different distances from the coast; their study did not detect any significant spatial distribution pattern¹³.

There is a generally held belief that the risk of contracting RRv disease is not related to place of residence but is randomly distributed in space. This is at least in part because the mosquitoes which transmit the disease in Brisbane are widely dispersed from their breeding sites by the prevailing onshore (easterly to north easterly) winds during the main breeding season (November to March). *Aedes vigilax* has been found as far as 50 km west of its main breeding areas in the coastal salt marshes¹⁴.

The main objective of this study was to assess the validity of the belief. The research questions were first, whether there are any areas within Brisbane which have significantly high or low infection rates and, second, if so, whether there is any spatial relationship with vegetation types, as indicators of mosquito breeding habitats.

Material and Methods

Study area

Brisbane, the state capital of Queensland, is situated on the east coast of Australia. Fig. 1 shows the location of the city at the mouth of the Brisbane River, which flows into Moreton Bay. The coastal areas are very flat, with extensive mangrove forests, salt marshes and mudflats. West of the densely populated city center the terrain is more hilly and to a large extent covered by rainforests and eucalypt forests. Within the administrative boundaries of Brisbane City Council, which also determine the study area of this investigation, the population size is approximately 800,000.

Disease data

As RRv is a notifiable disease, positive test results have, by law, to be reported by laboratories to the Queensland Department of Health, where they are archived by the Communicable Diseases Unit. The requirement for notification of RRv disease is based on a demonstration of IgM antibodies in blood, demonstration of a fourfold or greater change in serum antibody titres between acute and convalescent phase sera, isolation of RRv or demonstration of arboviral antigen or genome in blood. The common method used



Fig. 1 Location of the study area-Brisbane city centre is indicated by the arrow.

is the demonstration of IgM antibodies. Virus isolation is seldom used. Serological surveillance is not generally used in Queensland, and so data from such a source is not available. Whilst it is a legal requirement that laboratories report all RRv positive results, there are likely to be subclinical cases that are not noted. Issues of data reliability were addressed by Russell 15). There is no reason why there should be any significant spatial bias in the location of reported cases and so we are confident that the spatial distribution reflects the complete pattern, even if the numbers are underreported. Data from laboratory tests for the years 1991 to 1996 have been used. It is evident that the residential addresses of RRv patients need not necessarily be identical with the place where they contracted the virus. The infection could occur at work or during leisure time in any other location within or outside the study area. Nevertheless it is likely that, if there are differences found in the infection rate between suburbs, they are at least partly caused by different infection risks in those suburbs.

The data set includes information about the patients' address as well as sex and age and the onset date of disease. The address consists of house numbers, street and suburb names, postcodes and an additional code which indicates if this address refers to a private home or a hospital. For this study only private home addresses have been included. All non residential addresses such as hospitals and post office box numbers as well as those which could not be geocoded correctly have been excluded. Out of a total of approximately 3500 cases reported for Brisbane in the study period 2,252 records could be used.

Data preparation and geocoding

In order to secure privacy rights, the house numbers had to be randomised before the transfer from the Department of Health. Thus the house numbers used in this study differ from the original addresses by up to ± 10 . The records were geocoded in a GIS program. During this process, addresses from the database file were matched with a digital road network containing street data as well as suburb boundaries (ERSIS StreetInfo Brisbane) so that the addresses could be displayed as point data on a map.

In the first run only 1,597 out of the 2,252 address records (71%) could be processed successfully, ie. both the street name and the suburb name formed a valid combination. Apart from mistyping of street names in the original file the most frequent reason for this high error rate was incorrect or incomplete naming of suburbs which meant that addresses could not be immediately located. These various mismatches had to be corrected interactively in a second run; finally, 2,160 case records (96%) were successfully geocoded.

Demographic data

The spatial analysis was performed at the suburb level; Brisbane consists of 162 such units. The point data were overlaid with the suburb polygons and the cases per suburb were counted; for the demographic analysis the population size of each suburb was calculated as the mean value between the 1991 and 1996 Census records. The infection rate was standardised by the age distribution of the infected population. This was necessary as RRv infections occur primarily in adults. In some fast growing suburbs at the periphery of Brisbane the age distribution deviates significantly from the overall average with a higher proportion of children and youths, which would give an unrealistically low infection rate.

 Table 1
 Vegetation types investigated for correlation with RRv infection

Category	Description	Relevance for	Area	% of total
		mosquitoes	(km²)	study area
Littoral wetlands	Mangroves, saltmarshes	Breeding habitat for saltmarsh mosquito <i>Aedes</i> <i>vigilax</i>	21.0	2.18
Ephemeral wetlands	Areas temporarily covered with water after heavy rainfalls in summer and/or during extreme tides	Breeding habitat for saltmarsh mosquito <i>Aedes</i> <i>vigilax</i>	22.1	2.28
Open freshwater	Freshwater bodies with areas of aquatic vegetation	Breeding habitat for freshwater species	1.7	0.2
Riparian vegetation	Open forests or scrubland along banks of rivers and creeks	Breeding habitat for freshwater species	40.8	4.23
Melaleuca open forests	Open forests mainly Melaleuca spp. usually associated with wetlands	Breeding and resting habitat for freshwater species	18.4	1.91
Wet eucalypt open forests	Mixed open forests with wetland affinity	Breeding and resting habitat for freshwater species	22.0	2.28
Other bushland	Various other bushland types not directly associated with wetlands	Resting habitat	199.2	20.63

Rainfall data

Because differences between years may be affected by local weather, we obtained monthly rainfall data from the Australian Bureau of Meteorology, for the Brisbane Airport, as a rough indicator of rainfall in the study area.

Vegetation data

Brisbane City Council has an extensive environmental GISdatabase with detailed information on vegetation communities. For this study, seven vegetation types relevant as mosquito habitats have been selected for closer analysis (Table 1). In total, these types cover approximately 34% of the study area.

Statistical Analysis 1: Initial Analyses

Disease distribution. The hypothesis was whether there are significant differences between suburbs in the infection of RRv. The spatial distribution of RRv was tested for statistical significance using signed chi-square statistics: $\chi^{2} _{sig} = [(O - E)^{2}/E] * sig(O - E)$, where O is the observed number of cases per suburb and E the expected value, ie. the number of cases expected for each suburb as the product of the mean infection rate and the population size ¹⁶). The ranges used to map positive or negative deviations from the expected value have been chosen according to the 25, 5 and 1 percent significance levels.

Rainfall. To investigate if there were temporal relationships with rainfall, a simple linear regression was performed on age standardised infection rates and rainfall, by month and offset by 1, 2 and 3 months prior to disease notification, to take into account time elapsed between rainfall event, mosquito emergence and disease transmission and appearance of symptoms.

Statistical Analysis 2: Multivariate Analysis

Analysis of disease infection rate and vegetation types. Multivariate analyses were performed to discover relationships between vegetation types and their pattern in Brisbane related to RRv cases ¹⁷). Software packages involved in this procedure were MapInfo (GIS base module), ERSIS GIS Toolbox (GIS overlay functions), and SPSS (statistical analyses).

The distribution of the seven selected major vegetation types was intersected with the suburb layer. Percentage of coverage of each of the above vegetation types in each of the 162 suburbs in Brisbane was computed. A Principal Component Factor Analysis was performed in order to obtain a smaller number of condensed vegetation type factors that would explain most of the variability in the distribution of vegetation patterns in the study area. Factors were identified and ordered according to their importance.

The factors were then related to the RRv infection rate at the suburb level. The scores were computed and used as exploratory variables in a multiple regression to relate the disease infection rate to the vegetation factors ¹⁸. Factors that were significantly related to RRv rate were identified and ranked.

Results

Spatial Disease Distribution

There were significant differences between the suburbs in the infection rate of RRv disease as shown in Fig. 2 a and b. Fig. 2a shows the age-standardised distribution of the disease; Fig. 2b shows the significance of the differences between the suburbs.





aters

Indooroopilly Island

Oxley Creek

wetlands

The areas with the highest infection rates mostly coincide with known major mosquito breeding sites: the suburbs in the north-east of the study areas are situated at very low elevations, the terrain is characterised by creeks and coastal swamps and extensively covered with saltmarshes and mangrove forests. These ecosystems are permanent breeding sites of saltmarsh mosquitoes. There are also two areas of high infection rates along the freshwater systems of Oxley Creek in the southwest and in the densely vegetated areas around Tingalpa Creek in the southeast.

Another major area of high infection rates is in the western suburbs associated with the Brisbane Forest. These areas, although generally drier than the coastal wetlands, offer a favourable habitat for freshwater mosquitoes, mostly because of the dense vegetation cover and a large number of ephemeral creeks.

The city centre itself shows a relatively low infection rate. Although close to the Brisbane River, many inner city suburbs are located at a relatively high elevation, as the river itself is deeply incised. The river banks are very often unvegetated and do not offer much potential for mosquito breeding. A major exception to this can be seen west from the city centre: Indooroopilly Island is an important saltmarsh mosquito breeding site. South of the city centre there are many suburbs with relatively low infection rates, they are mostly located on the dry and hilly areas to the south of the city.

The results clearly indicate that the spatial distribution of RRv cases within the Brisbane area is not uniform and that there is apparently a relationship between the infection risk and major wetland areas.

Temporal Disease Distribution and Correlation with Rainfall

Because the data comprised disease infection rate aggregated over several years it was of interest to discern if individual years showed different patterns. There were significant differences between years in infection rates and these were significantly related to monthly rainfall. The most significant correlation was between the monthly rainfall and the number of cases two months later ($r^2 = 0.1259$, P = 0.0022). The pattern for the geocoded data is shown in Fig. 3. There are two reasons for the time lag. One relates to the natural cycle: following rainfall culicine disease vector mosquitoes must lay eggs on water, the eggs must hatch, larvae must develop to adulthood, adults must bite infected intermediate hosts, the virus must replicate in the mosquito which then bites and infects a person who later develops symptoms, visits a medical practitioner, has a blood sample taken and analysed. Then an estimated date of disease onset is provided for the data base. Secondly, data was provided on a monthly basis and so an early month notification might



relate back to a late month rainfall event in the penultimate preceding month.

Analysis of the data for individual years did not show any significant differences in the spatial patterns between years; the centroids of all years were in the center of the study area. The mean distance from the coastline was constant over all years, although there were minor differences in the N-S distribution parallel to the coastline.

Vegetation Factors and Their Relationship with RRv Distribution

Principal component factor analysis of the vegetation cover within suburbs yielded five factors. Table 2 shows the coefficients for each vegetation type as well as the proportion of variability explained by each factor. As can be seen from Table 2 the first five factors explained over 85% of the variance in the overall vegetation patterns. Factor 1 is a wetland factor, comprising ephemeral and littoral wetlands; factor 2 relates to wetlandassociated bushland (wet eucalypt and melaleuca forests), and factor 3 is mainly determined by riparian wetland. Of the remaining factors, factor 4 is a freshwater wetland factor, while factor 5 is a pure bushland factor without contribution of other wetland categories.

Linear regression analysis showed a high significance between RRv infection rate and factors 1, 3 and 5 (Table 3). These three factors represent three different land cover complexes: Littoral and ephemeral wetlands (factor 1), riparian wetlands (factor 3) and bushland (factor 5). The spatial distribution of these three types in the study area is illustrated in Fig. 4. The factor with the highest loading in the regression model was the bushland factor (factor 5), followed by the riparian vegetation (factor 3) and the wetlands (factor 1).

Discussion

The study showed that risk of RRv disease is spatially associated with certain vegetation types, and particularly with those which are obviously related to mosquito breeding such as wetlands. The model as described here is an important step towards a generally applicable RRv risk model. However, the method used here has some constraints which should be considered in interpreting the results and planning future research:

Table 2	Results of the	orincipal c	omnonent analysis -	vegetation factors
I adle Z	Results of the	Drincidai c	ombonent analysis -	vegetation factors.

Vegetation type	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Bushland (non-wetland)	04842	.01766	.04991	05219	.98296
Ephemeral wetlands	.83838	.03399	.20885	.21650	13293
Freshwater	.18849	01315	.06315	.96228	06044
Littoral wetlands	.89910	07242	17297	.00542	02175
Mixed open eucalypt	.00690	.95740	.06760	03352	00228
Melaleuca open forests	.56223	.39108	13996	.26473	.16150
Riparian wetlands	03851	.04607	.97588	.05408	.05044
% Variance explained	30.9	17.1	15.2	13.1	10.3

 Table 3
 Results of the linear regression model relating vegetation factors to RRv infection

Variable	Regression	STD Error	Standardised	t	Significance
	coefficient		coefficient		
Constant	331.250	20.346	16.261		
Wetlands (Factor 1)	41.952	20.409	0.148	2.056	0.041
Riparian (Factor 3)	69.111	20.409	0.244	3.386	0.001
Bushland (Factor 5)	88.150	20.409	0.311	4.319	0.000



Fig. 4 Map of the distribution of significant vegetation types (using MapInfo Australia)

Analysis at Suburb Level

The current model does not consider the location of the various land cover categories nor the distribution of the population within individual suburbs. This may be irrelevant for smaller suburbs with a diameter of just a few kilometres. However, in some larger outer suburbs the population is concentrated in a relatively small area. In these cases the proportion of certain vegetation cover categories may be very high, but the actual distance between the population center and these areas may also be very large. This is particularly true for the western suburbs of Brisbane which contain large uninhabited areas of rainforest, relatively far away from the major residential developments. In theory the use of smaller demographic units, such as the census collection districts (or census tracts), would help alleviate this problem. However in practice, at this level, the number of cases per unit was too small (2130 cases for 1541 districts) to determine any significant differences between the actual number of infections and the expected number calculated from the mean infection rate. It is planned for a future study to investigate the neighbourhood of each individual address and to compare the results with randomly chosen addresses.

Temporal Variability and Demographic Factors

The number of RRv infections varied significantly over the years. In the times series studied (1991-1996), peak years have been followed by years with relatively small numbers of infections. This could be due to different climatic conditions (see for example Fig. 3) but also due to some other epidemiological factors which have not yet been studied. For instance, after years of low RRv infection and/or immigration of susceptible people, a large outbreak of disease is likely (Weinstein)¹⁹. Of particular interest would be the question as to how long infected persons had lived at their current residential addresses. Weinstein suggested that high infection rates occur when non-immune populations move into endemic areas due to urban expansion. With the data set of our study, this hypothesis could not be supported: the mean age-standardised infection rate of 197 geocoded cases per 100,000 inhabitants for the 20 suburbs with the highest growth rate between 1986 and 1996 was even smaller

than the Brisbane average of 277.

Another issue may be that, in years with smaller numbers of infections, the disease would be centered around its natural focus area or nidus, which in this case would be the permanent mosquito breeding sites around the major wetland areas. It could be assumed that in peak years the virus would be spread over a larger area via the ephemeral breeding sites. The present analysis did not support this, with distributions not differing significantly between years. However, it is problematic to compare peak years to non-peak years statistically, as the number of cases in non-peak years is very low and therefore possible differences in the spatial distribution very difficult to verify.

Residential Address and Place of Infection

A patient's residential address need not necessarily coincide with the location where an infection occurred. The virus can also be contracted during weekend or holiday activities away from home as well as in the workplace, particularly for outdoor occupations such as gardeners or construction workers. This would be very difficult to assess. Currently, Brisbane City Council is conducting a major interview survey of RRv patients in order to detect correlations between their lifestyle and the disease, and these data could possibly offer some clues for further investigations.

Effects from Neighbouring Areas

In smaller suburbs there could be an influence from natural habitats in adjacent suburbs. A major wetland in one suburb can affect the mosquito density in a neighbouring suburb. This can also be relevant with respect to riparian complexes: As boundaries between suburbs are quite often defined by creeks, the associated riparian vegetation and wetlands will influence the mosquito density on both sides of the border. To overcome this problem it is planned to consider distance from each individual to the significant vegetation types in an extension to this study.

Other Vegetation Types as Potential Mosquito Habitats

The environmental database includes natural or semi-natural vegetation complexes which are relevant as mosquito habitats. Apart from these, mosquitoes can also breed in or migrate to residential areas with an appropriate habitat structure, such as gardens and non-natural parklands with dense vegetation cover or ponds. The general greenness of a suburb is likely to also be an important factor for the RRv infection risk. As there are no data available for this factor, a correlation could only be calculated using vegetation density indices or the Normalised Difference Vegetation Index (NDVI). This is a measure of vegetation based on the relationship between reflectance in the visible red end of the spectrum and in the near infrared (beyond the visible range) derived from remote sensing data.

Future Directions

The major task for the next step in this research project will be to systematically study the natural features in the vicinity of each geocoded address in order to identify dependencies between the probability of an infection and distances to or sizes of neighboring mosquito habitats as well as interactions between these various factors. This investigation will be conducted on the basis of digital land cover and vegetation data provided by the Brisbane City Council.

Acknowledgments

This study was supported by the Griffith University Research Grants Scheme. The authors want to thank Mr. Neil Davidson, Environmental Management Systems Unit, Brisbane City Council, for the provision of the environmental data, Mr. John

References

- Kitron U, Pener H, Costin C, Orshan L, Greenberg Z, Shalom U. Geographic information system in malaria surveillance: mosquito breeding and imported cases in Israel. Am J Trop Med Hyg 1994; 50: 550-6.
- Hugh-Jones, M. Introductory remarks on the application of remote sensing and geographic information systems to epidemiology and disease control. Prev Vet Med 1991; 11: 159-61.
- 3) Freier JE. Eastern Equine Encephalitis in Florida: the use of Geographic Information Systems. Paper presented at the 59th annual meeting of the American Mosquito Control Assoc. April 1993.
- Mott KE, Nuttall I, Desjeux P, Cattand P. New geographical approaches to control of some parasitic zoonoses. Bull World Health Organ 1995; 73: 247-57.
- 5) Kitron U, Michael J, Swanson J, Haramis L. Spatial analysis of the distribution of LaCrosse encephalitis in Illinois, using a geographic information system and local and global spatial statistics. Am J Trop Med Hyg 1997; 57: 469-75.
- 6) Arambulo III, PV, Astudillo V. Perspectives on the application of remote sensing and geographic information system to disease control and health management. Prev Vet Med 1991; 11: 345-52.
- 7) Beck LR, Rodriguez MH, Dister SW, et al. Remote sensing as a landscape epidemiologic tool to identify villages at high risk for malaria transmission. Am J Trop Med Hyg 1994 ; 51: 271-80.
- 8) Bogs F, Newell J, Fitzgerald J. Modeling spatial affects of landscape pattern on the spread of fungal disease in simulated agricultural landscapes. Proc.3rd Int.Conf.on Integrating GIS and Environmental Modeling, Nat.Centre for Geographic

Sheridan, Communicable Diseases Unit, Queensland Department of Health for the provision of RRv infection data and Mr George Hapgood of the Queensland Department of Health for further information in response to reviewers' comments.

Information Analysis, Santa Fe (CD-ROM) 1996.

- 9) Curran M, Harvey B, Crerar S, et al. Australia's notifiable diseases status. 1996 annual report of the National Notifiable Diseases Surveillance system. Communicable Disease Intelligence (Australia) 1997; 21: 281-307.
- Russell RC. Mosquitoes and mosquito-borne disease in southeastern Australia.Department of Medical Entomology, Westmead Hospital, Westmead, NSW 2145 1993.
- 11) Pavlovsky EN. Natural nidality of transmissable diseases. Univ of Illinois Press, Urbana 1966.
- 12) Barnes P, Dale PER, Muhar A. Landscape epidemiology, risk and mosquito control, Arb Res Aust 1997; 7: 8-11.
- 13) Mottram P, Smith D, Gould D, McGinn D, Sheridan J. Mosquito Prevalence in Selected Suburbs Locating at Different Distance from Salt Marsh Breeding Sites. Paper presented at the Seventh Arbovirus Resarch Symposium, Surfers Paradise, Queensland, Australia 1996.
- 14) Marks, EN. An atlas of common Queensland mosquitoes. Queensland Institute of Medical Research, Herston 1982.
- 15) Russell R.C. Ross River virus: disease trends and vector ecology in Australia. Bull Soc Vector Ecol 1994; 19: 73-81.
- Visvalingam M. The signed chi-square measure for mapping. Cartogr J 1978; 15: 93-8.
- 17) Krzanowski, WJ. Principles of Multivariate Analysis, A user's perspective. Oxford Science Publications, Oxford 1988.
- Draper N, Smith H. Applied Linear Regression (2nd Edn) John Wiley, New York 1981.
- 19) Weinstein P. An ecological approach to public health intervention: Ross River virus in Australia. Environ Health Perspectives 1997; 105: 364-6.