PC Lai 賴寶珍
CM Wong 黃浙明
AJ Hedley 賀達理
GM Leung 梁卓偉

Key Messages

1. Geographic information system (GIS) can be applied during an acute infectious disease outbreak to reveal new geospatial information in addition to standard field epidemiological analyses.
2. When applied in real time during the onset and evolution of an epidemic, GIS can monitor and enhance understanding of the transmission dynamics of an infectious agent, thereby facilitating the design, implementation and evaluation of potential intervention strategies.

Department of Geography, The University of Hong Kong, Pokfulam, Hong Kong SAR, China
PC Lai
Department of Community Medicine, The University of Hong Kong, 21 Sassoon Road, Hong Kong SAR, China
CM Wong, AJ Hedley, GM Leung

RESEARCH FUND FOR THE CONTROL OF INFECTIOUS DISEASES

Spatial clustering of SARS in Hong Kong

Introduction

Since the SARS outbreak, progress has been made in understanding the biology, pathogenesis and epidemiology of both the disease and coronavirus. Nonetheless, much remains to be done in the development of effective therapeutic interventions and diagnostic tools of sufficient sensitivity and specificity soon after the onset of symptoms. The evaluation of key epidemiological parameters and the impact of outbreak interventions at the population level remain to be investigated. Moreover, there was evidence of clear geographic micro-clusters of SARS cases and superspreading events (SSEs) which yielded widely varying densities of infection among the districts.[1]

The application of geographic information system (GIS) in health and health care is a new approach. In particular, a wide variety of cartographic techniques for mapping and analysis of communicable disease data have become available. These enable the application of GIS to examine spatially related problems from different perspectives. In addition to descriptive mapping functions, GIS is capable of data manipulation and geostatistical analysis.

Aims and objectives

The GIS technology was applied to map and visualise the SARS outbreak in Hong Kong. Cartographic and geostatistical methods were used to represent and analyse patterns of disease spread during the 2003 outbreak. This study also assessed the utility and limitations of GIS as a real-time disease surveillance tool.

Methods

Both spatial and non-spatial data were used. Spatial data are geographic and represent the real world features as points, lines, or areas. SARS occurrences were depicted as points on a base map of Hong Kong. Descriptive data about the confirmed cases of SARS were derived from case-contact interviews. Associated residential address data were first cleaned and then verified before undergoing geo-referencing to enable mapping.

We analysed the SARSID integrated database which contained details on all confirmed SARS cases admitted to hospitals in Hong Kong during this epidemic (from 15 February 2003 to 22 June 2003). A total of 1709 confirmed cases (out of 1755 cases) were extracted for geo-coding and analysis. The unaccounted cases (2.6% of the total) had inconsistencies in the address entries.

Three levels of analysis were carried out: (1) elementary, (2) cluster, and (3) contextual. Simple visual inspection of the SARS incidence was conducted at the elementary level, followed by an identification of hot spots through cluster analysis. At the contextual level, the relationships between SARS and various geographical phenomena were examined. All analyses were carried out using the ArcGIS and its extension modules by the Environmental Systems Research Institute in Redlands, California.

Results

Maps generated in elementary analysis revealed the geographic spread of SARS

infection by residential addresses of Hong Kong. There was clear clustering of cases in certain districts of the Kowloon peninsula (such as Kwun Tong) and the New Territories (including Shatin and Tai Po), but Hong Kong Island was relatively spared.

Cluster analysis generated statistical surfaces using the kernel density method to account for SARS incidents on the date of symptom onset and a 5-day incubation period. Each of the time-series kernel maps showed SARS infection rate per 1000 inhabitants on a prototypical day during the 16-week epidemic, with darker zones emphasising disease hot spots.

Contextual analysis enabled construction of daily histograms of the number of observations by 15 classes of infection rates, which were primarily composed of inverse J-shaped curves, showing an increased concentration of SARS occurrences toward the end of March 2003. Origin-and-destination plots of disease clusters were used to explore likely or probable locations of index cases or environmental sources of infection as informed through contact tracing by public health authorities.

Discussion

The kernel method of portraying infection rates provided a means of highlighting locations of disease risks. The related R-values and Moran's coefficients enhanced the analytical context of the point-pattern distributions. Such geospatial intelligence provided the basis for formulating our transmission dynamics model.[2] A variety of approaches ranging from a simple deterministic compartmental approach to a spatially explicit and individual-based simulation were possible in constructing the transmission dynamics model. We based our analyses on a stochastic meta-population compartmental model because the incidence of SARS varied substantially according to geographical districts.

The daily animated series of kernel maps clearly show that SARS was a highly localised disease. In contrast with influenza and measles transmitted through casual contact, the route of transmission for SARS was more compatible with close contact via heavy respiratory droplets and fomites. An alternative interpretation of the observed high degree of geospatial clustering shows that SARS was attributed to an environmental point source outbreak, as demonstrated by faulty sewage systems and the chimney effect hypothesised for the Amoy Gardens SSE. Although it is difficult to gauge retrospectively, had the GIS been available for near real-time analysis, it might have afforded more rapid contact tracing and public health interventions to prevent further large-scale environmental point source outbreaks.

Contextual analysis is a useful adjunct to the usual bio-mathematical modelling approach using reproductive numbers at different points in time throughout the SARS epidemic.[2] The origin-and-destination analysis complemented with R and Moran's I values suggested the direction of spread in a disease cluster that could be used to inform contact tracing and the design of quarantine measures. Instead of isolating entire residential districts as practised in China at the height of the SARS outbreak, these analytical approaches might have enabled better selection of such districts for quarantine.

There are limitations to the GIS technique in infectious disease epidemiology and outbreak investigation. Mapping of diseases tends to expose the 'where' but not 'why there' of the outbreak although map patterns can provide stimuli for generating hypotheses of disease causation. Moreover, newer developments that complement traditional mapping functions such as cluster and contextual analyses can be useful adjunctive investigational tools in outbreak control.

The completeness and rapid availability of necessary data is another area of concern. Conventional field epidemiological data rarely contain the full range of variables required in a GIS analysis. Unfortunately, non-standardisation of patient address formats and missing details diminish the proportion of useable cases for analyses. A number of generic problems associated with information system development must be resolved to render real-time disease monitoring and surveillance. On top of the list is standardisation of data capture documents, as well as procedures and protocols for information management. There is also an urgent need to manage delays in transferring and updating disease information to facilitate rapid analysis and audit of databases. The SARS epidemic is a clear signal that Hong Kong needs much greater investment in health informatics (ie public health information systems, the skills to use them, and networks to share them).

Conclusions

Integration of GIS technology into routine field epidemiological surveillance offers a scientifically rigorous and quantitative method for the identification of unusual disease patterns in real time. Its potential can be synergistically maximised when linked to clinical databases collecting data at the point of care. This integration should entail the whole population and environmental data sources (including meteorological, transportation, topographical information) so as to rapidly recognise, locate and monitor disease outbreaks.

Acknowledgements

References

1. Donnelly CA, Ghani AC, Leung GM, et al. Epidemiological determinant of spread of causal agent of severe acute respiratory syndrome in Hong Kong. Lancet 2003;361:1761-6.

2. Riley S, Fraser C, Donnelly CA, et al. Transmission dynamics of the etiological agent of SARS in Hong Kong: impact of public health interventions. Science 2003;300:1961-6.