

Clusters of autochthonous hepatitis A cases in a low endemicity area

H. GUIIS¹, S. CLERC¹, B. HOEN² AND J. F. VIEL^{1*}

¹ *Department of Public Health, Biostatistics and Epidemiology Unit, Faculty of Medicine, Besançon, France*

² *Department of Infectious Diseases and Tropical Medicine, University Medical Center and Faculty of Medicine, Besançon, France*

(Accepted 2 August 2005, first published online 5 October 2005)

SUMMARY

At the University Hospital of Besançon (département of Doubs, France), an unusually high number of patients were hospitalized for hepatitis A during the 1999–2000 period, some of whom had not travelled abroad. This prompted us to conduct an investigation on a population basis and search for clusters of cases possibly related to local sources of contamination. Accordingly, case definition was restricted to autochthonous cases. During the 1999–2002 period, 45 autochthonous cases were classified as possibly originating from local environmental sources. A space–time scan statistic detected one most likely cluster (standardized incidence ratio 20·63, 95% confidence interval 10·6–37·1), consisting of 11 persons (of whom five children had attended the same swimming pool). It remained significant in a sensitivity analysis, strongly supporting the hypothesis of an environmental source of contamination. This study reveals the necessity of regular surveillance for hepatitis A and raises the issue of virological surveys of pool waters.

INTRODUCTION

Hepatitis A virus (HAV) is transmitted by the faecal–oral route, directly by person-to-person contact or indirectly by ingestion of raw or inadequately cooked contaminated food or water. Direct transmission occurs most frequently within households or closed communities such as day-care centres or other institutions [1]. Shellfish, which can concentrate the virus, or raw vegetables washed with infected water, are the most common sources of contaminated food [2–6]. The virus is unusually stable and can resist several months in contaminated soils, marine sediment, fresh water and seawater [1].

Worldwide distribution of hepatitis A is heterogeneous [7]. High levels of endemicity patterns of HAV infection are encountered in developing countries with very poor sanitary and hygienic conditions (parts of Africa, Asia and Central and South America), where infection is usually acquired during early childhood. Intermediate levels are found in developing countries, countries with transitional economies and some regions of industrialized countries where sanitary conditions are variable (Southern and Eastern Europe, some regions in the Middle East). As in most developed countries, the incidence of hepatitis A is low in France and the disease occurs among specific risk groups such as travellers returning from endemic areas.

In France, the national notifiable disease surveillance system does not encompass the notification of hepatitis A. Epidemiological data, crucial for formulating strategies and policies for the prevention

* Author for correspondence: Professor J. F. Viel, Department of Public Health, Biostatistics and Epidemiology Unit, Faculty of Medicine, 2 place Saint Jacques, 25030 Besançon, France.
(Email: jean-francois.viel@ufc-chu.univ-fcomte.fr)

and control of this disease are, therefore, dependent on the willingness and cooperation of medical physicians, microbiological laboratories, and public health practitioners.

At the University Hospital of Besançon (département of Doubs, France), an unusually high number of patients with hepatitis A were hospitalized during the 1999–2000 period, some of whom had not travelled abroad in the 2 months preceding their illness. This prompted us to set up an investigation to identify the hepatitis A cases on a population basis and search for clusters of cases possibly related to local sources of contamination.

POPULATION AND METHODS

Study area

The département of Doubs (5233 km², 499 162 inhabitants in 1999) is located in eastern France. It is divided into 594 'communes' (the most precise spatial unit for which population data were available) whose populations range from 6 to 117 550 inhabitants.

Case definition

A case was defined by a positive serology for IgM anti-HAV between 1 January 1999 and 31 December 2002, in patients living in the département of Doubs. Excluded were patients who had an IgM hyperproduction (e.g. myeloma), or had been vaccinated against HAV in the year before the serology.

To further select the autochthonous cases, patients were excluded if they had travelled to a HAV endemic country [7] 2 months prior to the first day of illness.

Finally, we attempted to differentiate autochthonous patients possibly infected by a local environmental source of contamination (LESC) from autochthonous patients possibly contaminated by an index case. Thus, in a conservative approach, patients who had been in contact during the 2 months preceding their symptoms with a person presenting with hepatitis A and returning from an endemic area, were excluded. We assumed that these patients were more likely to have been contaminated through direct contact with an index case than by a local environmental source.

Case identification

Cases were collected retrospectively in 2003. Information sources consisted of laboratory records,

hospital records and minimum in-patient datasets. Eighteen (72%) of the 25 laboratories of the département of Doubs agreed to participate in the study. The other seven declined to participate because they had no computerized record databases.

Information regarding the date of the serology as well as the name and address of the patient and his/her general practitioner (GP) were collected. GPs were then contacted and asked to obtain from their patient an informed consent for participation in the study. Patients who were unable to be contacted by their GP were contacted directly by mail.

Ethical clearance for this study was granted by the Consulting Committee for the Treatment of Information in Medical Research (no. 02.394) and by the National Commission for the Confidentiality of Computerized Data (no. 903079).

Demographic data

Population data by commune, gender and 5-year age group were obtained from the French Office of Population Censuses for the 1999 census. In the cluster detection method, expected numbers of cases for each commune were computed by applying an internal standard (i.e. incidence rates from the whole département for the same years) to the number of person-years for each area, stratified by gender and 5-year age groups (indirect standardization).

Questionnaire

Once the patient agreed to participate, a phone interview was arranged to complete a questionnaire assessing the circumstances of HAV contamination. Similar to the one used by the French National Institute for Public Health Surveillance [8], the questionnaire contained items related to three topics: (1) identification of the patient and demographic characteristics; (2) clinical symptoms and laboratory markers; and (3) risk factors for HAV contamination (trips to endemic areas, contacts with HAV-infected people, contacts with or visits to facilities caring for infants and young children or people living in closed communities, exposure to contaminated water, exposure to contaminated food, blood transfusion and use of intravenous drugs).

Cluster detection

To test for clusters of cases and to locate them, we used the space–time scan statistic developed by

Kulldorff et al. [9, 10]. We chose to geocode the commune of the home address during the 2-month period preceding the symptoms. Considering that the incubation period is quite variable (7–45 days), instead of choosing the exact date of the serology, we chose the month as time unit. We assumed that the numbers of cases were Poisson distributed.

A cylindrical window that continuously changed in centre, radius and height scanned the geographic area for potential clusters. More precisely, the window moved from the centroid of one commune to the centroid of another commune. For each location, the radius varied continuously from zero to a maximum such that at most 30% of the département population was included (when the window included a centroid, the population relating to the commune was included as a whole in the window), and the height of the cylinder (representing the time dimension), varied from zero to a maximum so that the window never included more than 30% of the 4-year period. For each centroid, cylindrical windows, therefore, included different sets of neighbouring communes and time periods.

For each location and size of the scanning window, the null hypothesis was that the incidence rate was the same for each commune. Conversely, under the alternative hypothesis, the incidence rate was assumed higher inside than outside the window. During the space–time process, an enormous amount of different cylinders were evaluated to find the most likely cluster. A likelihood function was calculated for each cylindrical window and was maximized over all of them. The most likely space–time cluster was the one with the maximum likelihood, corresponding to a given location, radius and time-frame. Its *P* value was obtained through Monte Carlo hypothesis testing (9999 replications). A result was considered significant at the 5% level.

RESULTS

Cases

From 1999 to 2002, 95 patients living in the département of Doubs tested positive for IgM anti-HAV antibodies. Hepatitis A was confirmed by physician diagnosis for 87 cases (one patient presented a myeloma, and seven were lost to follow-up). We excluded 42 people because these cases were travel related. The remaining 45 patients were considered as possibly contaminated by a LESC.

Incidence rates

The incidence rate was 4·36/100 000 when considering all confirmed hepatitis diagnoses (87 cases).

When focusing on LESC cases, the incidence rate was 2·25/100 000 (45 cases). Higher rates were observed in persons aged <15 years (4·39/100 000), and persons aged >85 years (4·97/100 000). Cases originated from 19 of the 594 communes and annual incidence by commune varied from 0 to 208 cases per 100 000.

Searching for space–time clusters

Two non-overlapping and significant space–time clusters were detected (Fig., Table 1). The most likely one consisted of 34 communes in the western part of the département. In this area, 11 cases occurred in November–December 2000 when 0·53 were expected ($P=0\cdot0001$). The secondary cluster corresponded to a unique commune in the northeastern part and included four cases diagnosed in December 2000–January 2001 when 0·05 were expected ($P=0\cdot0157$).

Of the 15 cases who comprised the two clusters, nine belonged to four families. Since the virus is excreted 1 or 2 weeks before the onset of symptoms and the incubation period can sometimes last only 1 week (even if in most cases it usually lasts 1 month), it may not be possible to differentiate a common environmental contamination (since members of a family often share the same activities) from a secondary person-to-person contamination within the family. Thus, in a sensitivity analysis, we kept only the first person diagnosed in each family (37 cases instead of 45). The same most likely cluster was detected whereas the secondary cluster was not (Table 2). The analysis was carried out two more times to include the seven lost to follow-up: one time with 52 cases (45 LESC+7), the other time with 44 cases (37+7). Results remained globally similar (Table 2).

To obtain a more complete picture on HAV transmission routes, we have also searched for clusters of travel-related hepatitis A cases. A highly significant cluster encompassing 23 communes (of which one corresponded to the secondary cluster of LESC cases) was found in the northeastern part of the département. It consisted of 13 cases that occurred between August 2000 and January 2001 when 0·95 were expected ($P=0\cdot0001$). No secondary cluster of travel-related cases was detected.

Table 1. Space-time clusters of autochthonous hepatitis A cases possibly originating from a local environmental source of contamination* in the département of Doubs, France, 1999–2002

Clusters	No. of communes	Time period	Observed cases (expected cases)	SIR (95% CI)	P value
Most likely	34	November to December 2000	11 (0.53)	20.6 (10.6–37.1)	0.0001
Secondary	1	December 2000 to January 2001	4 (0.05)	83.2 (21.8–204.8)	0.0157

SIR, Standardized incidence ratio; CI, confidence interval.

* These cases corresponded to patients with a positive serology for IgM anti-hepatitis A virus (HAV) of whom patients presenting a myeloma, or having travelled to a HAV-endemic country were excluded.

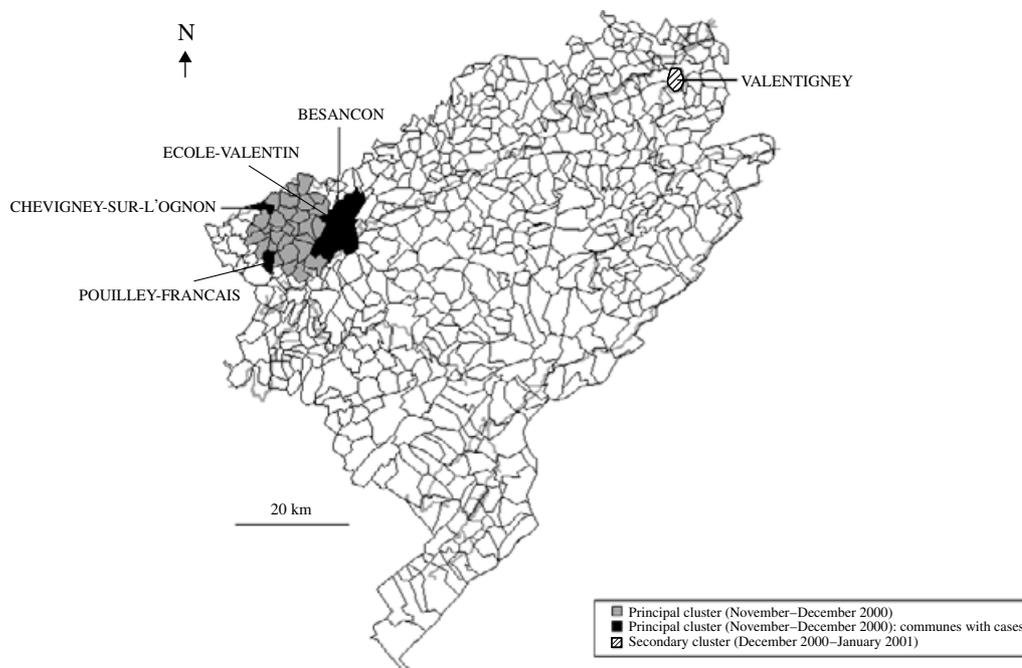


Fig. Space-time clusters of hepatitis A cases, possibly originating from a local environmental source of contamination (département of Doubs, France, 1999–2002).

Searching for an environmental source at the origin of the clusters

We explored the possible contamination sources encountered by the cases comprising the clusters during the 2-month period preceding their symptoms.

The most likely cluster of LESC cases included eight children (aged 3–15 years), and three young adults (aged 23–34 years). Six were females, five were males. One adult belonged to a group considered at risk (health-care provider working with children). The two pre-school children attended different day-care facilities. One case had eaten shellfish during the 2 months preceding his symptoms, another one

reported that he might also have done so. Two cases sometimes ate food brought back by their family from abroad (dates and pastries from Algeria and fresh tropical fruits) but could not remember whether they had done so during the 2-month period. Five children had gone to the same public swimming pool during the 2-month period preceding their symptoms.

The secondary cluster of LESC cases included only children ($n=4$), three of whom belonged to the same family. The second and third children had been diagnosed with hepatitis A 1 month after the diagnosis in the first sibling. The fourth child had been in contact with a friend who had developed hepatitis A in the 2-month period. No other possible

Table 2. Sensitivity analysis according to case definition (space–time clusters of autochthonous hepatitis A cases possibly originating from a local environmental source of contamination* in the département of Doubs, France, 1999–2002)

Case definition	Cluster type	Time period	Observed cases (expected cases)	SIR (95% CI)	P value
First cases of each family (37 cases)	Most likely	December 2000	8 (0.44)	18.0 (7.8–35.8)	0.0014
Cases and lost to follow-up (52 cases)	Most likely	November to December 2000	11 (0.61)	18.0 (9.0–32.2)	0.0001
	Secondary	December 2000	6 (0.05)	107.2 (44.0–261.2)	0.0001
First cases of each family and lost to follow-up (44 cases)	Most likely	September 2000 to January 2001	12 (1.38)	8.7 (4.4–15.2)	0.0016
	Secondary	December 2000 to April 2001	5 (0.12)	43.4 (13.5–97.2)	0.0129

SIR, Standardized incidence ratio; CI, confidence interval.

* These cases corresponded to patients with a positive serology for IgM anti-hepatitis A virus (HAV) of whom patients presenting a myeloma, or having travelled to a HAV-endemic country were excluded.

contamination source was evoked apart from one case who had eaten shellfish.

DISCUSSION

Case definition

By design, this study is limited to people with a symptomatic hepatitis which led them to consult a physician and to undergo a serological test, therefore excluding asymptomatic hepatitis A. When corrected for asymptomatic infections and under-reporting, one case of hepatitis A notified to the surveillance system would represent an estimated 10 infections (as reported in the United States, a low endemicity area) [11].

Case definition aimed at restricting the study to autochthonous cases possibly originating from a LESC but, as a result, impeded comparisons with LESC incidence rates from other areas, since no other sources mentioning rates of autochthonous hepatitis A were found in the literature.

Completeness

Laboratory-based reporting of serological markers for viral hepatitis is an increasingly common route by which suspected cases are identified to health authorities, increasing the completeness and timeliness of case identification [11]. Surprisingly, 28% of the laboratories could not retrieve retrospective

information due to the absence of a computerized record database. Unfortunately, we did not have a second independent source of data to estimate the number of missing cases by a capture–recapture method. In a pilot laboratory-based surveillance study, carried out in nine French départements, the participation of the laboratories varied from 35% to 75% depending on the département [12]. In the present study, the participation rate, therefore, ranks high.

Three of the seven laboratories that did not participate were geographically close. Thus, we cannot rule out that the probability of detecting clusters from that area might be reduced or that the two clusters identified might be less significant. However, the identification of two cases in this area and its low-density population made an under-reporting bias less likely. Furthermore, this potential bias is only spatial, not temporal, so its influence on the significance of space–time clusters should be only minor, if at all.

Finally, the overall hepatitis incidence rate observed in this study (4.36/100 000) is similar to those observed in five of the nine départements involved in the French pilot study (3–5/100 000) [12], giving further support to the completeness of our study.

Incidence

Incidence varied greatly across communes. Highest incidence rates should be interpreted with caution

since the communes with such rates are also the communes with the smallest populations and, thus, the less stable estimates.

LESC incidence peaked in November–December. Symptomatic or asymptomatic cases returning from vacations in countries where HAV is endemic might be responsible for secondary hepatitis A, by direct contact or by contaminating environmental sources. It would have been interesting to compare strains of HAV to check whether the same strains are involved in contamination of cases returning from areas where HAV is endemic and cases contaminated later on in winter (clustering on strains or subtypes of viruses would be in favour of a unique common source) [13–15]. Unfortunately, investigating the genetic variability of HAV strains (by sequencing selected genome regions) is still an area of research, currently beyond the expertise of many microbiological laboratories [16].

Incidence of LESC cases was high in two groups: children, and more surprisingly, people aged >85 years. High incidence in the elderly can be explained by two factors [17]. First, the increased risk of morbidity and mortality when infected with HAV at this age entails a more prompt serological testing by physicians. Second, the elderly are at risk for hospitalization (whatever the cause), and in hospitals, serological tests are prescribed more often than in medical practices. Therefore, they might have had a greater chance of being detected.

Clusters

We assumed that if an environmental source of HAV existed, a space–time cluster should be detected in the vicinity. HAV can resist several months in the environment, so an environmental source can be at the origin of contaminations during a few months, or during a longer period if the source is contaminated repeatedly. We, therefore, searched for clusters during a time period of 1–16 months (30% of the time period at most). A smaller time unit might have suffered from variations of incubation period.

When geocoding home addresses, we assumed that exposition occurred in the commune of residence. The cluster could stretch, at most on 30% of the département surface, allowing a search for clusters on a larger area than the commune level. The bias due to migration of people might exist but would result

in a dilution effect, reducing chances of detecting clusters, but not challenging their validity.

Kulldorff's method was chosen because it has several advantages: it adjusts for population density and confounding variables (e.g. age, gender); there is no preselection bias since the clusters are searched with no prior hypothesis on their location, size or time period; the test statistic takes into account multiple testing and delivers a single *P* value; if a cluster is detected, its location and time-frame is specified [18].

However, our approach has two limitations. First, it was impossible to adjust for other possible confounding factors (such as socio-economic levels or levels of education) [1, 19], since these data were not available for the commune/gender/age group combinations. Second, this method is more powerful for detecting clusters with a compact shape, than more elongated ones, such as clusters that stretch along a river.

Two significant clusters were detected. The most likely cluster, located around Besançon, and lasting from November to December 2000 was very robust since it persisted when cases that could have been contaminated by family contact were excluded and when cases lost to follow-up were included. This consistency strengthens the reliability of our results. This cluster strongly supports the hypothesis of an environmental source of contamination. Furthermore, the briefness of the cluster is in favour of this hypothesis since epidemics due to a common source are usually short whereas those due to person-to-person contact can last for lengthy periods (e.g. 18 months) [20].

The secondary significant cluster spreading from December 2000 to January 2001, disappeared when cases that could have been contaminated by family contact were excluded. It was yet again detected when cases lost to follow-up were included, whether cases who could have been contaminated by family contact were excluded or not. This cluster might represent the extension of the travel-related case cluster highlighted in the same area. Indeed, all persons but one composing these clusters belonged to the large community of North African descent living in this region (Algeria was the country of origin for most of immigrants in the 1960s and early 1970, when the economic expansion, mainly mass car production in this area, generated an enormous need for workers). The time overlap (travel-related cases: August 2000–January 2001;

LESC cases: December 2000–January 2001) suggests a contamination of LESC cases by direct contact with people having travelled to North Africa during the summer.

Sources of contamination of cases composing the most likely cluster

For the most likely cluster, the only environmental source suggested was a public swimming pool located in Besançon. Other studies have revealed that pools could be at the origin of hepatitis A contaminations [21, 22]. No virological survey of water quality of public pools is performed routinely in France. Results of surveys concerning total coliforms, faecal coliforms, streptococcus and free residual chlorine were available from the city council, and showed that water quality fulfilled standard guidelines from August 2000 to January 2001.

Nevertheless, HAV has previously been isolated in dam-water samples where indicators of faecal pollution were absent or at low levels [23]. Furthermore, HAV seems to be more resistant to chlorine than other enteric viruses: it is not inactivated by concentrations of free residual chlorine recommended in public pools (0.4–1.4 mg/l) [24, 25]. This study confirms the necessity of investigating more thoroughly HAV inactivation conditions in pools. It raises the issue of the virological survey of pool water and highlights the importance of reassessing quality guidelines based on indicator organisms.

Overall, this study reveals the necessity to have in place a system for regular time periodic disease surveillance to detect any currently active geographical clusters of hepatitis A, or other infectious diseases [26].

ACKNOWLEDGEMENTS

We thank all the laboratories, physicians and patients who agreed to participate in the study. This research was supported by the Health and Rural Environment Research Group (EA 2276), University of Franche-Comté, France. We are indebted to an anonymous reviewer whose comments led to improvements in this paper.

DECLARATION OF INTEREST

None.

REFERENCES

1. **Fiore AE.** Hepatitis A transmitted by food. *Clin Infect Dis* 2004; **38**: 705–715.
2. **Chironna M, Germinario C, De Medici D, et al.** Detection of hepatitis A virus in mussels from different sources marketed in Puglia region (South Italy). *Int J Food Microbiol* 2002; **75**: 11–18.
3. **Dentiger CM, Bower WA, Nainan OV, et al.** An outbreak of hepatitis A associated with green onions. *J Infect Dis* 2001; **183**: 1273–1276.
4. **Mullendore JL, Sobsey MD, Carol Shieh Y.** Improved method for the recovery of hepatitis A virus from oysters. *J Virol Methods* 2001; **95**: 25–35.
5. **Romalde JL, Torrado I, Ribao C, Barja JL.** Global market: shellfish imports as a source of reemerging food-borne hepatitis A virus infections in Spain. *Int Microbiol* 2001; **4**: 223–226.
6. **Wang JY, Hu SL, Liu HY, Hong YL, Cao SZ, Wu LF.** Risk factor analysis of an epidemic of hepatitis A in a factory in Shanghai. *Int J Epidemiol* 1990; **19**: 435–438.
7. **Centers for Disease Control and Prevention.** Geographic distribution of hepatitis A prevalence, 2000 (http://www.cdc.gov/travel/diseases/maps/hav_map.htm). Accessed 28 February 2005.
8. **Delarocque-Astagneau E, Hemery C, Duchon C.** An outbreak of hepatitis A: Midi-Pyrénées, France [in French], 1997 (<http://www.invs.sante.fr/publications/default.htm>). Accessed 28 February 2005.
9. **Kulldorff M, Rand K, Williams G.** SaTScan, version 3.1, Software for the spatial and space-time scan statistic. Bethesda, MD: National Cancer Institute, 1997 (<http://www.satscan.org>). Accessed 28 February 2005.
10. **Kulldorff M, Athas WF, Feuer EJ, Miller BA, Key CR.** Evaluating cluster alarms: a space-time statistic and brain cancer in Los Alamos, New Mexico. *Am J Public Health* 1998; **88**: 1377–1380.
11. **Centers for Diseases Control and Prevention.** Guidelines for viral hepatitis surveillance and case management. USA, Atlanta, 2005.
12. **Couturier E, Delarocque-Astagneau E, Vaillant V, Desendos JC.** Hepatitis A surveillance in France during the last two decades: current data do not permit estimations of incidence rates [in French]. *Bulletin Epidémiologique Hebdomadaire*, 2005 (http://www.invs.sante.fr/beh/2005/05/beh_05_2005.pdf). Accessed 28 February 2005.
13. **Bruisten SM, van Steenberg JE, Pijl AS, Niesters HG, van Doornum GJ, Coutinho RA.** Molecular epidemiology of hepatitis A virus in Amsterdam, the Netherlands. *J Med Virol* 2001; **63**: 88–95.
14. **Tjon GMS, Wijkman CJ, Coutinho RA, et al.** Molecular epidemiology of hepatitis A in Noord-Brabant, the Netherlands. *J Clin Virol* 2005; **32**: 128–136.
15. **van Steenberg JE, Tjon GMS, van den Hoek A, Koek A, Coutinho RA, Bruisten S.** Two years' prospective collection of molecular and epidemiological data shows limited spread of hepatitis A virus outside risk groups in Amsterdam, 2000–2002. *J Infect Dis* 2004; **189**: 471–482.

16. **Costa-Mattioli M, Cristina J, Romero H, et al.** Molecular evolution of hepatitis A virus: a new classification based on the complete VP1 protein. *J Virol* 2002; **76**: 9516–9525.
17. **Brown GR, Persley K.** Hepatitis A epidemic in the elderly. *South Med J* 2002; **95**: 826–833.
18. **Kulldorff M.** Statistical evaluation of disease cluster alarms. In: Lawson A, Biggeri A, Böhning D, Lesaffre E, Viel JF, Bertollini R, eds. *Disease mapping and risk assessment for public health*. Chichester, England: John Wiley & Sons, 1999; 144–149.
19. **Fix AD, Martin OS, Gallicchio L, Vial PA, Lagos R.** Age-specific prevalence of antibodies to Hepatitis A in Santiago, Chile: risk factors and shift in age infection among children and young adults. *Am J Trop Med Hyg* 2002; **66**: 628–632.
20. **Begue P, Bernuau J, Courouce AM, et al.** Prevention of hepatitis A virus transmission during epidemics [in French]. *Bulletin Epidémiologique Hebdomadaire*, 1996 (<http://www.invs.sante.fr/beh/>). Accessed 28 February 2005.
21. **Mahoney FJ, Farley TA, Kelso KY, Wilson SA, Horan JM, McFarland LM.** An outbreak of hepatitis A associated with swimming in a public pool. *J Infect Dis* 1992; **165**: 613–618.
22. **Hubert B, Bourderieux C, Dubois M, Marquis M.** An outbreak of hepatitis A associated with the frequenting of a paddling pool [in French]. *Bulletin Epidémiologique Hebdomadaire*, 1999 (<http://www.invs.sante.fr/beh/1999/9904/index.html>). Accessed 28 February 2005.
23. **Taylor MB, Cox N, Vrey MA, Grabow WOK.** The occurrence of hepatitis A and astroviruses in selected river and dam waters in South Africa. *Water Res* 2001; **35**: 2653–2660.
24. **Li JW, Xin ZT, Wang XW, Zheng JL, Chao FH.** Mechanisms of inactivation of hepatitis A virus by chlorine. *Appl Environ Microbiol* 2002; **68**: 4951–4955.
25. **Peterson DA, Hurley TR, Hoff JC, Wolfe LG.** Effect of chlorine treatment on infectivity of hepatitis A virus. *Appl Environ Microbiol* 1983; **45**: 223–227.
26. **Kulldorff M.** Prospective time periodic geographical disease disease surveillance using a scan statistic. *J Royal Statist Soc (A)* 2001; **164**: 61–72.