Identifying high-risk areas for sporadic measles outbreaks: lessons from South Africa

Benn Sartorius, C Cohen, T Chirwa, G Ntshoe, A Purenb & K Hofmana

Abstract

Introduction

Measles is highly contagious and the transmission intensity of the measles virus exceeds that of most human pathogens. Despite progress in controlling measles since 2000, the disease remains endemic in many countries and killed over 164 000 children worldwide in 2008. Deaths also occur during outbreaks in areas where the disease is no longer endemic.

Measles vaccination is highly effective, safe and relatively cost-effective and has interrupted measles transmission in most parts of the world, although high-risk areas remain in Africa and southern Asia. Immunization coverage is a key indicator for monitoring health sector performance and progress towards Millennium Development Goal 4 (i.e. reducing child mortality). measles is difficult to control and eliminate and more than 90% of the population must be immune to interrupt transmission and prevent outbreaks. Maintaining adequate vaccine coverage throughout countries is a global problem and, in 2010 and 2011, outbreaks occurred even in the developed world. Moreover, since the efficacy of the first vaccine dose is thought to be around 85%, non-responders will add to the pool of susceptible children. In addition, achieving a high level of population immunity may be difficult in areas such as sub-Saharan Africa where the prevalence of human immunodeficiency virus (HIV) infection is high because infected children are less likely to respond to vaccination or maintain a protective antibody level or may acquire lower antibody levels from an HIV-positive (HIV+) mother. Vaccination may need to be repeated more frequently in these areas.

Since 1995, several African countries have launched initiatives to eliminate measles, as recommended by the World Health Organization (WHO). These initiatives aim to: (i) achieve a coverage of 90% or more for the first vaccine dose; (ii) identify areas with low coverage and high-risk areas for "mop-up" immunization of children who missed vaccination; and (iii) enhance case-based measles surveillance with laboratory support.

In South Africa, first and second measles vaccine doses have been given free of charge to children aged 9 and 18 months since April 2009 as part of the WHO Expanded Program on Immunization. At present, measles vaccination is mandatory and provided in the public sector, whereas combination vaccination against measles, mumps and rubella is available in the private sector at a cost. The South African National Department of Health introduced case-based measles surveillance in 1998 with laboratory support from the National Institute for Communicable Diseases.

A large outbreak of measles involving around 1700 cases occurred in South Africa between 2003 and 2005 following its introduction from Mozambique. More recently, there was an outbreak between 2009 and 2011, with over 18 000 cases recorded. The transient trough in cases following a large outbreak, which increases herd immunity, affords planners an opportunity to investigate the epidemic and refine prevention strategies. Spatial analysis has become an important epidemiological tool for detecting and predicting patterns of disease spread. In situations in which population-wide interventions are too expensive or are inefficient because they dilute available resources, the effectiveness of vaccination is increased by targeting high-risk areas, such as those implicated in measles transmission and those where vaccination coverage varies.

The objectives of this study were: (i) to investigate the pattern of measles spread in the South African outbreak
between 2009 and 2011; (ii) to identify associations between the outbreak and population density, HIV prevalence and preceding lapses in measles vaccination; and (iii) to develop a model for identifying areas at a high risk for future large-scale sporadic outbreaks that can be used in prevention.

Methods

Administratively, South Africa is divided into 9 provinces and 52 districts, which include 6 metropolitan areas. Districts are further disaggregated into 248 local municipalities, with metropolitan areas being governed by metropolitan municipalities. Health service delivery is largely a municipal function.

We obtained data on individual cases observed between 2008 and 2011 from the measles surveillance programme at the National Department of Health and the National Institute for Communicable Diseases. Data on measles vaccination coverage among infants (i.e. children aged under 1 year), on the proportion of HIV+ infants aged under 2 months and on relevant denominators (e.g. the total number of infants in a given district, by year) were obtained from the District Health Information System. The main unit of analysis was the district.

We calculated the overall incidence of measles in each week between 2009 and 2011 to identify temporal peaks. In addition, we calculated the incidence in infants in each district by dividing the observed number of infant measles cases by the total infant population in the district. To identify districts in which the incidence was significantly above the average, either before or during the outbreak, we derived 95% confidence intervals (CIs) for each incidence proportion using the binomial distribution of the observed number of events.

The incidence in a district was considered significantly above average in a particular year if the national incidence for that period was below the lower limit (i.e. $\alpha = 0.025$) of the incidence proportion for the district.

Modelling the risk of measles

Standardized risk ratios (RR) were calculated for each district as the ratio of the observed to the expected number of infant measles cases in each district. The expected number of cases was the national measles incidence multiplied by the infant population of the district.

The most widely used strategy for tackling problems posed by small-area analysis is Bayesian hierarchical modelling. The specific approach we adopted was developed by Besag, York and Mollié and includes two random-effects terms (i.e. the convolution approach) that account for: (i) structured district
contiguity and (ii) unstructured district heterogeneity. It is the most widely used spatial Poisson model and the basic kriging formulation, which includes no covariates, is as follows:

\[ O_i \sim \text{Poisson}(E_i \lambda_i) \]

\[ \log(\lambda_i) = \alpha + \beta X_i + \epsilon_i + \omega_i \] (1)

where \( O_i \) is the observed number of measles cases in district \( i \), \( E_i \) is the expected number of cases, \( \lambda_i \) is the Poisson mean or expected value (and the variance), \( \alpha \) is a constant, \( \epsilon_i \) is the unstructured district heterogeneity random effect and \( \omega_i \) is the conditional autoregressive spatial random effect accounting for district contiguity. Further details are provided in Appendix A (available at: http://www.fileswap.com/docs/BTMr0X3gU/Appendix1.doc).

This basic spatial model can also be extended to a space–time formulation: the full formulations involving Bayesian inference using Gibbs sampling (BUGS) are provided in Appendix B (available at: http://www.fileswap.com/docs/dl/fbTMr0X3gU/Appendix2.doc). If a large number of areas have excess or repeat zero counts, which is often the case with rare outcomes, then a zero-inflated Poisson approach must be used. 26–28

In anomaly detection, Bayesian exceedance probabilities were used to detect areas in which the risk was significantly increased. 29 We plotted model estimates on district maps that depicted smoothed estimates of the RR and the posterior probability that the RR was greater than 1. We used a more stringent version of Richardson’s criterion, 29 in which a district with a Bayesian exceedance probability (i.e. the probability that the RR would exceed 1) greater than 0.9 was deemed to be a high-risk area.

To investigate the influence of preceding lapses in vaccination, the prevalence of HIV infection among infants and vaccine failure on the observed risk of measles, we tested various determinants at the bivariate and multivariate level using the Bayesian approach described above with:

\[ \log(\lambda_i) = \alpha + \beta X_i + \epsilon_i + \omega_i \] (2)

where \( X_i \) represents the vector of the associated covariates and \( \beta \) represents the corresponding regression coefficients. Factors found to be significant at the bivariate level (i.e. \( P < 0.2 \)) were included in the multivariate model. Districts were divided into tertiles according to population density: low, medium and high density. All metropolitan areas were classified as high density but there were many high-density areas that were not metropolitan areas. We created a dummy variable with three categories to reflect this: low- to medium-population-density area, high-population-density nonmetropolitan area and high-population-density metropolitan area.

The formula below was developed to use data from the preceding year to estimate the size of the infant popula-
tion susceptible to measles in each district, which is possibly the most important predictor of future high-risk areas. In this study, the proportion of infants who were not covered by vaccination in a given year was defined as the proportion that did not receive the first measles vaccine dose in that year. The proportion that was not immunized was the combination of the proportion not covered by vaccination and the proportion that would fail to seroconvert following the first vaccine dose, taking into account the influence of HIV infection on the likelihood of seroconversion. A previous study showed that mortality in HIV-infected children without access to antiretroviral therapy can limit the effect of the HIV epidemic on increasing measles virus transmission. Consequently, we refined the formula by factoring in expected mortality rates among HIV-negative (HIV−) and HIV+ infants, with the proportion of HIV+ infants being defined as the proportion receiving antiretroviral therapy. The interaction between the proportion susceptible to measles and population density was also assessed.

The size of the infant population susceptible to measles in each district in year $t$ (i.e. $S_t$) was derived using data from year $t-1$ for that district, by adding those not covered to those with vaccine failure among covered, as follows:

$$S_t = \left[ Pop_{t-1} \times (1 - P_{\text{dose}\ t-1}) \times P_{\text{HIV+}} \right]$$

where $Pop_{t-1}$ is the infant population of the district in year $t-1$, $P_{\text{dose}\ t-1}$ is the proportion of infants who receive their first measles dose in year $t-1$, $P_{\text{HIV+}}$ is the proportion of all infants who survive to the age of 1 year, $P_{\text{fail HIV+}}$ is the proportion of infants who are HIV+ in year $t-1$, $P_{\text{fail HIV+}}$ is the proportion of HIV+ infants who fail to seroconvert after the first measles vaccine dose, $P_{\text{on ART} t-1}$ is the proportion of HIV+ infants receiving antiretroviral therapy in year $t-1$, $P_{\text{HIV+}}$ is the

**Fig. 4. Elevated risk of measles in infants in districts during the first outbreak peak, a South Africa, 2009**

The shaded districts are those in which the risk was significantly elevated as indicated by their Bayesian exceedance probabilities determined using a Bayesian convolution conditional autoregressive model.

a The first outbreak peak occurred around week 43 in 2009.

**Fig. 5. Elevated risk of measles in infants in districts during the second outbreak peak, a South Africa, 2010**

The shaded districts are those in which the risk was significantly elevated as indicated by their Bayesian exceedance probabilities determined using a Bayesian convolution conditional autoregressive model.

a The second outbreak peak occurred around week 10 in 2010.
The values of some of the parameters used in Equation 3 were based on data from the District Health Information System for 2006 to 2011. The proportion of infants who failed to seroconvert after the first vaccine dose was obtained from a recent meta-analysis: 64% (95% CI: 49–78%) of HIV+ infants seroconverted by 9 months after measles vaccination, compared with between 62 and 100% (weighted average: 87%) of those who were exposed to HIV but did not become infected and between 76 and 92% (weighted average: 85%) of those who were not exposed to HIV. Hence, $P_{\text{fail HIV}+}$ was 0.36 (i.e. $1–0.64$) and $P_{\text{fail HIV}–}$ was 0.14 (i.e. $[1–0.87]+[1–0.85]$ divided by 2). Values for the proportion of infants who survived to the age of 1 year were taken from a pooled analysis of data from sub-Saharan Africa, in which the estimate for $P_{1,\text{HIV}+}$ was 0.648 and, for $P_{1,\text{HIV}–}$, 0.951. The value for $P_{1,\text{all}}$ was derived from a weighted estimate of the proportion of infants who were HIV+ or HIV– and their respective expected survival proportions. The proportion of HIV+ infants receiving antiretroviral therapy was estimated to be 0.19 in 2008, 0.27 in 2009, 0.31 in 2010 and 0.40 in 2011.

The analysis was carried out using Stata (StataCorp. LP, College Station, United States of America) and WinBUGS (MRC Biostatistics Unit, Cambridge, United Kingdom of Great Britain and Northern Ireland). Risk maps were constructed in MapInfo Professional (Pitney Bowes Inc., Stamford, USA).

**Results**

Demographic data showed that, during the 2009 to 2011 measles outbreak, the incidence of the disease was highest among infants: 61 per 10,000 (95% CI: 59.3–62.4). The second highest incidence was in children aged 1 to 4 years: 7.3 cases per 10,000 (95% CI: 7.0–7.6). The incidence in children aged 10 to 14 years and in those aged 15 to 19 years was higher than in those aged 5 to 9 years. The baseline incidence among infants before and after the outbreak was 1.0 per 10,000 (95% CI: 0.9–1.3).

There were two peaks in measles incidence in infants during the outbreak: the smaller peak occurred in week 43 of 2009 and the slightly larger peak occurred approximately 20 weeks later, in week 10 of 2010 (Fig. 1). Spatial analysis identified districts in which the risk

---

**Fig. 6.** Predicted relative risk of measles in infants in a populated area, by proportion of susceptible infants

---

**Fig. 7.** Relative risk of measles in infants, by district population density, South Africa, 2009–2011

---

The circles indicate the observed relative risk and population density in individual districts. The solid line represents the fitted association between the observed relative risk and population density and the dashed line represents the lower bound of the 95% confidence interval for the association.
of measles was significantly increased over the course of the outbreak (Fig. 2 and Fig. 3). The distribution of disease incidence was distinctly nonrandom (Moran’s I global \( P < 0.001 \)). Several neighbouring (i.e. connected) high-risk districts were observed: in Western Cape Province in the south-west of the country, in Gauteng and Mpumalanga Provinces in the north and in KwaZulu-Natal Province on the east coast (Fig. 2). In addition, the incidence was elevated in metropolitan areas. Fig. 3 shows the estimated incidence of measles in infants in all districts, the national baseline incidence in infants before the start of the outbreak and the national incidence during the outbreak. The two districts with the highest incidence were in Northern Cape and KwaZulu-Natal Provinces respectively, followed by the Johannesburg (Gauteng) and eThekwini (KwaZulu-Natal) metropolitan areas.

The first significant elevation in risk occurred in the Tshwane metropolitan area during weeks 35 and 36 of 2009, before the first peak in incidence. By week 37, the significant rise in risk had spread to the neighbouring metropolitan areas of Johannesburg and Ekuruleni. An analysis of the period leading up to the first peak suggested that the outbreak was concentrated in Gauteng province. Fig. 4 and Fig. 5 show the districts most affected by the outbreak at the two peaks in incidence. The districts with the highest incidence were similar at the two peaks, though the outbreak was more widespread at the first peak. There appeared to be no spread into surrounding districts.

### Factors associated with measles outbreaks

The average coverage of the first dose of measles vaccine in all districts was 88.1% (range: 85.4–90.7). Many districts repeatedly failed to meet the target of 90%: 56.9%, 49.0% and 64.5% of districts had coverage below this critical value in 2008, 2009 and 2010, respectively. The risk of measles in infants in a district showed a significant positive association with the proportion of unvaccinated infants and with the proportion of infants not immunized in the preceding year (Table 1). In addition, the risk of infection was increased in metropolitan districts, all of which had a high population density (\( P < 0.001 \)). The model predicted that the critical cut-off proportion for coverage with the first vaccine dose is 0.2. Consequently, the risk would be elevated (i.e. RR > 1) in highly populated areas in which the percentage of susceptible infants exceeds 20% because of poor vaccination coverage in preceding years (Fig. 6). Bivariate interactions indicate that the threshold would be higher in less densely populated areas. Districts with a population density of more than 750 individuals per square kilometre also appear to be at an elevated risk (Fig. 7). Using data on the level of coverage of the first vaccine dose
in districts in 2011, the model predicted that the high-risk areas in which the proportion of susceptible infants would exceed the critical cut-off level in 2012 lay mainly in KwaZulu-Natal (Fig. 8). Since these areas are highly connected (i.e. neighbouring), the outbreak could spread rapidly.

Discussion

The study showed that HIV infection, preceding lapses in vaccination and high population density were associated with an increased risk of measles outbreak and the predictive model we developed provides a means of identifying areas at high risk of a future outbreak or where herd immunity does not meet elimination criteria. In resource-limited settings, it is essential to focus on high-risk areas where vaccination coverage is poor and where the build-up of a susceptible population could fuel future outbreaks. Our model, which takes into account the influence of HIV infection, could be applied to other settings where sporadic measles outbreaks remain a problem and can be modified easily for application to other vaccine-preventable diseases.

Infants were the group at highest risk in the measles outbreak we examined. This is not surprising given that their immune systems are undeveloped. Although maternal measles antibodies generally protect infants in the first 6 months of life, studies suggest that the infants of mothers with vaccine-induced immunity and of HIV+ mothers receive significantly fewer antibodies. In addition, HIV infection in infants in sub-Saharan Africa has been shown to reduce their ability to seroconvert following vaccination. These factors have implications for herd immunity: the build-up of a susceptible population in an area with low vaccine coverage and a high prevalence of HIV infection could lead to a large measles outbreak. In areas where measles and HIV infection are prevalent, both HIV-infected infants and infants exposed to HIV but not infected could benefit from earlier vaccination.

We found that, overall, national coverage with the first measles vaccine dose in South Africa was high, at around 88%, but was variable, with the coverage in many districts routinely falling well below 90%. A previous study in the country concluded that the heterogeneity in measles vaccine coverage challenges the goal of eliminating measles and that supplemental vaccination activities have a negative effect on routine immunization coverage and health system functioning. Areas with repeatedly low vaccine coverage should be targeted for improvement. It is also important to coordinate supplemental immunization activities between neighbouring countries during widespread outbreaks. It has been reported that the cases observed immediately following supplemental immunization tend to be due to immigration from nearby areas where it has not taken place.

Our findings suggest that both the build-up of a susceptible population due to poor vaccine coverage and high population density contributed to the South African outbreak. Similarly, the large measles outbreak in Sierra Leone during 2009 and 2010 was probably caused by the accumulation of susceptible individuals, largely due to nonvaccination. The cumulative impact of nonvaccination is important and justifies using our predictive model to help identify and target areas with a large number of susceptible individuals. We observed that the measles outbreak in South Africa had a strong nonrandom spatial distribution with strong connectivity, as seen in other studies. Previous studies also indicate that measles outbreaks generally persist through epidemic troughs when the community is above a critical size. Below this threshold, measles tends towards extinction during the troughs. Our findings indicate that once the percentage of the population susceptible to measles in a densely populated district exceeds 20%, the risk of infection is increased. Districts with a smaller susceptible population appeared remarkably resilient to measles infiltration during the outbreak. Our observation that the risk of measles outbreak was significantly elevated in areas with a high population density and in metropolitan areas has also been reported in other settings. The highly connected nature of these areas may have contributed to the persistence of the outbreak.

Study limitations include possible underreporting of measles cases, fragmentation of vaccination services, a change in measles vaccination schedule in 2009 and repeated changes in
reporting forms during the past 5 years. Underreporting in particular districts may have affected the spatial pattern observed and weakened associations with key indicators. In addition, we cannot discount the possibility of ecological bias given that aggregated measles vaccination and HIV infection data were used.

Our study has important implications for the elimination of measles in South Africa and sub-Saharan Africa. Lapses in immunization, a high burden of HIV infection and the build-up of a susceptible population appear to be important for driving sporadic outbreaks, especially in populous areas. Supplemental immunization is by definition a catch-up strategy. It is also costly, may not be implemented quickly enough to limit outbreaks and has a negative effect on routine immunization services. Identifying and targeting emerging high-risk areas in resource-limited settings where vaccine coverage is low or wanting appears a more viable strategy for preventing outbreaks in sub-Saharan Africa than mass supplemental immunization.

Acknowledgements
We thank Professor John Pettifor. Cheryl Cohen is also affiliated with the School of Public Health, University of the Witwatersrand.

Competing interests: None declared.

Abstract
Identifying high-risk areas for sporadic measles outbreaks in South Africa

Objective
To develop a model to identify regions at risk of sporadic measles outbreaks in South Africa, based on the spatial distribution of reported cases during 2009 and 2011.

Methods
Geospatial data on reported cases and vaccination coverage were obtained from the De Villiers acronym-based(Exceptional) Influenza Sentinel Surveillance System (D-VISS), a surveillance system for influenza and other respiratory infections in South Africa. The cases were linked to the National Immunization Programme (NIP) database to determine vaccination status and immunization history. The model was developed using a Bayesian hierarchical Poisson regression model, with vaccination coverage, population density, and HIV prevalence as covariates. The model was validated using a leave-one-out cross-validation technique.

Results
The model predicted a high risk of sporadic measles outbreaks in specific regions of South Africa, particularly in theWestern Cape and Eastern Cape provinces. The model also identified regions with low vaccination coverage and high HIV prevalence as being at higher risk of outbreaks.

Conclusion
The model developed in this study can be used to identify high-risk areas for sporadic measles outbreaks in South Africa and other similar settings. Further studies are needed to validate the model in other regions and to evaluate its effectiveness in preventing outbreaks.

Resume
Identifying les zones à haut risque d’épidémies sporadiques de rougeole: les leçons de l’Afrique du Sud

Objectif
Développer un modèle permettant d’identifier les zones à haut risque d’épidémies sporadiques de rougeole, sur la base d’une analyse des facteurs associés à une épidémie nationale en Afrique du Sud entre 2009 et 2011.

Méthodes
Les données sur les cas survenus avant et pendant l’épidémie nationale ont été obtenues auprès du programme sud-africain de surveillance de la rougeole, et les données sur la vaccination contre la rougeole et la taille de la population auprès des systèmes d’informations sanitaires des districts. Un modèle hiérarchique bayésien de Poisson a été utilisé pour étudier l’association entre le risque de rougeole chez les nourrissons dans un district et la couverture vaccinale pour la première dose, la densité de population, la prévalence sous-jacente de l’infection.
Результаты определения новых районов повышенного риска в 2012 году.

Модель Пуассона. Проекции модели использовались для сероконверсии использовалась байесовская иерархическая
иммунодефицита человека (ВИЧ) и ожидаемым отсутствием
населения; сопутствующим преобладанием инфекции вируса
у детей в округе с охватом первой дозы вакцины, плотностью
а также данные по противокоревой иммунизации и численности
корью, зафиксированных до и во время национальной вспышки,
них населения, сопутствующим преобладанием инфекции вируса
иммунодефицита человека (ВИЧ) и ожидаемым отсутствием
сериоконверсии использовалась байесовская иерархическая
модель Гуассона. Проекции модели использовались для
определения новых районов повышенного риска в 2012 году.

Результаты: Отмечена четкая пространственная картина
районов повышенного риска со многими взаимосвязанными
(т.е. соседними) районами. Повышенный риск возникновения
вспышек кори в значительной степени был связан как с
предшествующим ростом уязвимых слоев населения, так и с его
плотностью. Риск возникновения вспышек также повышался,
когда более 20% детей в густонаселенном районе пропускали
первую дозу вакцины. Данная модель позволила определить
районы повышенного риска возникновения вспышек кори в 2012
году, и районы, где могли быть предприняты профилактические
меры.

Вывод: Вспышка кори в Южной Африке была связана с ростом
уязвимых слоев населения (по причине недостаточного охвата
населения прививкой), высоким уровнем заболеваемости ВИЧ-
инфекций и высокой плотностью населения. Разработанная
модель прогнозирования могла бы применяться в других
условиях, подверженных спорадическим вспышкам кори и другим
инфекциам, контролируемым методами вакцинопрофилактики.

Resumen
Identificación de las áreas con riesgo alto de sufrir brotes esporádicos de sarampión: lecciones desde Sudáfrica

Objetivo: Desarrollar un modelo para identificar las áreas con riesgo alto de sufrir brotes esporádicos de sarampión en base a un análisis de los factores asociados a un brote nacional en Sudáfrica entre los años 2009 y 2011.

Métodos: Los datos de los casos ocurridos antes y durante el brote nacional se obtuvieron del programa de vigilancia del sarampión de Sudáfrica, y la información acerca de la vacunación contra el sarampión y el tamaño de la población, del sistema de información sanitaria por distritos. Se empleó un modelo de Poisson jerárquico bayesiano para investigar la relación entre el riesgo de sarampión en niños de un distrito con la cobertura de la primera dosis de la vacuna, la densidad de población, el historial de prevalencia de la infección por el virus de la inmunodeficiencia humana (VIH) y el fracaso previsto de la seroconversión. Las proyecciones del modelo se utilizaron para identificar áreas de alto riesgo nuevas en 2012.

Resultados: Se observó un patrón espacial claro de las áreas de alto riesgo con muchas zonas conectadas (esto es, colindantes). El riesgo superior de sufrir un brote de sarampión se asoció de modo significativo tanto con el crecimiento previo de un grupo de población vulnerable como con la densidad de población, y también fue elevado cuando más del 20% de los niños en un área populosa no había recibido la primera dosis de la vacuna. El modelo fue capaz de identificar las áreas con riesgo alto de sufrir un brote de sarampión en 2012 y donde podrían emprenderse medidas de prevención adicionales.

Conclusión: El brote de sarampión en Sudáfrica estuvo asociado con el aumento de un grupo de población vulnerable (debido a una cobertura de vacunación insuficiente), una prevalencia elevada de la infección por VIH y una densidad de población alta. El modelo predictivo desarrollado podría aplicarse a otros lugares susceptibles de sufrir brotes esporádicos de sarampión y otras enfermedades inmunoprevenibles.

References


