Estimating the measles effective reproduction number in Australia from routine notification data

May Chiew, Heather F Gidding, Aditi Dey, James Wood, Nicolee Martin, Stephanie Davis & Peter McIntyre

Objective To estimate the measles effective reproduction number (R) in Australia by modelling routinely collected notification data.

Methods R was estimated for 2009–2011 by means of three methods, using data from Australia’s National Notifiable Disease Surveillance System. Method 1 estimated R as 1 – F, where F equals the proportion of cases that were imported, as determined from data on place of acquisition. The other methods estimated R by fitting a subcritical branching process that modelled the spread of an infection with a given R to the observed distributions of outbreak sizes (method 2) and generations of spread (method 3). Stata version 12 was used for method 2 and Matlab version R2012 was used for method 3. For all methods, calculation of 95% confidence intervals (CIs) was performed using a normal approximation based on estimated standard errors.

Findings During 2009–2011, 367 notifiable measles cases occurred in Australia (mean annual rate: 5.5 cases per million population). Data were 100% complete for importation status but 77% complete for outbreak reference number. R was estimated as < 1 for all years and data types, with values of 0.65 (95% CI: 0.60–0.70) obtained by method 1, 0.64 (95% CI: 0.56–0.72) by method 2 and 0.47 (95% CI: 0.38–0.57) by method 3.

Conclusion The fact that consistent estimates of R were obtained from all three methods enhances confidence in the validity of these methods for determining R.

Abstract in العربية, 中文, Français, Русский and Español at the end of each article.

Introduction

Between 2000 and 2010, the global incidence of measles and the mortality caused by the disease decreased by 66% and 74%, respectively. The implementation of childhood and supplementary immunization activities by Member States of the World Health Organization (WHO) has largely contributed to these declines. As a result of this progress, target dates for measles elimination have been set by all WHO regions, most recently the South-East Asia Region. At present, the Region of the Americas is the only WHO region to have achieved elimination, with interruption of endemic measles transmission in 2002.

Three years after the declaration of measles elimination in the Americas, WHO’s Regional Office for the Western Pacific adopted a resolution that set 2012 as the target year for regional measles elimination. Several countries, including Australia, have made significant progress towards this goal.

Over the past two decades, the national incidence of measles in Australia declined dramatically following several national measles control strategies, including the 1998 Measles Control Campaign and revision of the age at which the second dose of measles, mumps and rubella vaccine is routinely administered (i.e. 4 years, down from 10–16 years). This led to a report that indigenous measles transmission had been eliminated in Australia, with the authors reasoning that Australia had satisfied most of the WHO regional office’s criteria for elimination of measles.

One approach to monitoring measles elimination involves estimating the effective reproduction number (R). R is the average number of secondary cases that result from an infectious case in a particular population (Box 1). R depends on the level of susceptibility in the population, in contrast to the basic reproduction number (R0), which is the average number of secondary cases arising from one infectious case in a totally susceptible population. When R is < 1, the number of cases decreases with every generation, and if this value is maintained elimination is considered to have occurred.

Several types of data have been used to estimate R, including serosurveillance data, the early epidemic phase growth rate and notification data. In Australia, three national serosurveys have been conducted. Serosurveys, however, can be time-consuming and costly, and they may not be feasible in developing countries. Alternatively, routinely collected notification data can be used to estimate R on a regular basis with minimal resources. These methods to estimate R were described by De Serres et al. and require the following data: the proportion of cases imported, the distribution of outbreak sizes and the distribution of the number of generations of spread in each outbreak. An assumption required when using these methods is that elimination has already been achieved.

Methods based on routinely collected notification data have been used to differing extents internationally and in Australia. Canada and the United States of America both reported the interruption of indigenous measles transmission by 1998 and used the methods described by De Serres et al. to assess elimination. England and Wales also claimed elimination and used data from 1995 to 2001 on the distribution of outbreak sizes to estimate R. In the Australian state of Victoria, R was estimated using notification data from 1998 to 2003 on the distribution of outbreak size (R = 0.87) and the generation of spread (R = 0.73). More recently, data...
on the proportion of imported cases were obtained from Australia's National Notifiable Diseases Surveillance System (NNDSS) to estimate R for 2001 to 2006 (R = 0.90) but the authors concluded that the incompleteness of the data may have resulted in the overestimation of R.9 There have been considerable improvements in the completeness of notification data since 2009. Therefore, the aim of our study was to use routinely collected notification data to estimate R and track measles elimination in Australia.

**Methods**

**Notification data**

In Australia, measles is a notifiable disease in all states and territories. Under public health legislation, clinicians and laboratories are required to notify their respective health authorities of suspected, probable or confirmed cases of measles.21 Additional information on confirmed cases is collected by health authorities during follow-up investigations and forwarded to the NNDSS. Confirmed cases require either laboratory-confirmed evidence of measles virus infection or both clinical and epidemiological evidence.21

De-identified data were obtained from the NNDSS data set for all notified measles cases between 2009 and 2011. Data fields used in the analysis included place (country) of acquisition, outbreak reference number (Box 1) and date of symptom onset. The date of diagnosis was used. A flowchart was developed to link cases for which data on the outbreak reference number were missing (Fig. 1).

If the place of acquisition was overseas, cases were considered sporadic unless they had the earliest onset date among cases in an identified outbreak and all other cases in the outbreak were locally acquired. If more than 1 case was acquired overseas in the same outbreak, the imported case most temporally similar to locally acquired cases was considered the index case, whereas the other imported case was considered sporadic.

**Estimating R**

Three methods were used to estimate R for 2009 to 2011 combined and for 2009, 2010 and 2011 individually. In method 1, R was estimated as 1 − P, where P is equal to the proportion of cases that were imported, as determined from data on place of acquisition. The derivation of this formula is available in the article by De Serres et al.;22 the formula assumes that all cases must be linked to an imported case and depletion of the pool of susceptible individuals can be ignored.

In the other two methods, R was estimated by fitting a subcritical branching process that modelled the spread

---

**Box 1. Definitions**

1. **Imported case**: a case public health authorities believe was acquired overseas based on international travel in the period before rash onset.

2. **Generation of outbreak spread**: 
   - Same generation: Disease onset date between the first and last case was 0–6 days apart.
   - One generation: Disease onset date between the first and last case was 7–14 days apart.
   - Two generations: Disease onset date between the first and last case was 15–24 days apart.
   - Additional generations: A 10-day difference between the disease onset date of the first and last case added an extra generation.

3. **Outbreak reference number**: a unique identification number assigned to cases that were determined to be part of the same outbreak on the basis of epidemiological (and/or virological) evidence.

4. **Effective reproduction number (R)**: 
   - R = 1: This is a state of endemic equilibrium in which, on average, one case results in one secondary infection.
   - R > 1: The number of cases increases from one generation to the next, potentially resulting in an epidemic.
   - R < 1: The number of cases decreases with each generation; if the decrease is maintained, elimination occurs.

---

**Fig. 1. Flowchart to determine links between measles cases for which outbreak reference numbers were missing**

- Missing outbreak reference number
- Same jurisdiction
- Symptom onset date
- Diagnosis date
- Genotype
- Residence within 35-km radius
- Linked
- Not linked
- Difference of 7–18 days
- Difference of <7 days or >18 days
- Missing
- Not linked
- Yes
- No
of an infection with a given $R$ onto the observed distributions of outbreak sizes (in method 2) and generations of spread (in method 3). Details of the branching process model are available elsewhere. Maximum likelihood estimation was used to obtain estimates of $R$ from the models. Stata, version 12 (StatCorp LP, College Station, USA), was used to estimate $R$ for the distribution of outbreak sizes, whereas estimates of $R$ using the distribution of the generations of spread were calculated in Matlab, version R2012 (The MathWorks Inc., Natick, USA).

All methods are based on the assumption that $R$ is < 1 and are only appropriate in settings where the measles incidence is low. In addition, for methods 2 and 3, it is assumed that the number of secondary cases arising from each infected individual is Poisson distributed with a mean $R$; thus, (non-random) heterogeneity between individuals is ignored. For all methods, 95% confidence intervals (CIs) were calculated using a normal approximation based on the estimated standard errors. Because larger outbreaks might be more likely than smaller outbreaks to be detected by the surveillance system, a sensitivity analysis was conducted for method 2 that included only outbreaks involving three or more cases.

Results

Notification data

Between 2009 and 2011, 367 cases of measles were notified in Australia (mean notification rate: 5.5 cases per million population); 105 in 2009 (notification rate: 4.8 cases per million population), 69 in 2010 (notification rate: 3.1 cases per million population) and 193 in 2011 (notification rate: 8.5 cases per million population). Most cases occurred in individuals aged 10 to 19 years (annual range: 30–38%) and 20 to 34 years (annual range: 34–41%). Thirty-five per cent of cases (128) were acquired overseas and 41% (150) were epidemiologically linked to an imported case. Data on place of acquisition and outbreak reference number were 100% complete (367 of 367 records) and 77% complete (283 of 367 records), respectively.

Eighty-four cases had no outbreak reference number and were initially considered sporadic. Of these, 19 (23%) were reclassified as belonging either to an identified outbreak or were linked together as part of an outbreak not identified previously, using the flowchart specified in Fig. 1. Seven cases were reclassified as belonging either to a cluster of two cases or a sporadic case, respectively.

Table 1. Estimation of the effective reproduction number ($R$) in Australia annually during 2009–2011

<table>
<thead>
<tr>
<th>Datum</th>
<th>$R$ (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2009</td>
</tr>
<tr>
<td>Proportion of imported cases</td>
<td>0.66 (0.57–0.75)</td>
</tr>
<tr>
<td>Distribution of outbreak sizes</td>
<td>0.65 (0.49–0.80)</td>
</tr>
<tr>
<td>Distribution of outbreak with at least 3 cases</td>
<td>0.78 (0.57–1.00)</td>
</tr>
<tr>
<td>Distribution of generations of spread</td>
<td>0.38 (0.19–0.56)</td>
</tr>
</tbody>
</table>

Discussion

From 2009 to 2011, $R$ estimates based on national notification data remained < 1, providing good evidence that measles elimination is being maintained in Australia. These findings are consistent with the $R$ of 0.69 that was estimated using the 2002 serosurvey data and with the prediction that, between 2003 and 2012, $R$ would remain < 0.80. Conclusion that measles elimination is being maintained is further supported by high measles vaccine coverage rates and the absence of an endemic circulating genotype for many years in Australia.

Similarities in $R$ estimates between the three methods used in this study and that of the serosurvey suggest that $R$ estimates based on Australian surveillance data are valid. Additionally, there was little difference between the estimates when restricting the outbreak size to 3 or more cases and when no restrictions were used, which indicates a high level of sensitivity of the surveillance system in detecting both small and large outbreaks.

There was, however, some variation among the estimates of $R$ across the three methods. The highest $R$ was observed in 2009, using the sensitivity analysis of model 2, and this was the
only value for which the upper limit of the CI reached 1.00. A large proportion of cases in 2009 were sporadic or occurred in clusters of two, leaving only seven outbreaks in the sensitivity analysis. Because of this small sample size and because the largest outbreak during the study period (25 cases) occurred during 2009, the sensitivity analysis involving method 2 had a wide CI and a high R estimate.

The R estimates were expected to be highest using method 1 owing to an underestimation of the number of cases arising from contact with international visitors, as was observed in similar studies conducted in Canada and the United States.11,12,13 In the United States the estimate of R based on importation status was 0.68 (95% CI: 0.30–0.78), compared with 0.51 (95% CI: 0.44–0.59) when using the distribution of outbreak size.13 Similarly, in Canada the estimate of R based on importation status was 0.87 (95% CI: 0.76–0.98), compared with 0.82 (95% CI: 0.72–0.93) when using the distribution of outbreak size.14 Our estimates, however, were comparable to the estimates obtained from the other two methods. The 100% ascertainment of data on place of acquisition is possibly the reason for our comparable results. In contrast, the completeness of the reporting of imported cases in the United States from 1997 to 2001 was only 40%.15 It is likely that the R values calculated using method 1 in our study provided accurate estimates for Australia because the surveillance system is sensitive and the data on place of acquisition were complete.

An important difference to consider when comparing our estimates with those from Canada and the United States relates to the definition of an imported case of measles. A case of measles in Canada is considered imported if exposure occurred outside Canada 7 to 21 days before rash onset and the case is not linked to local transmission.12 In the United States, travel outside the country within 18 days of rash onset – unless symptoms began at least 7 days after the date on which a travelling companion developed symptoms – is required for a case to be considered imported.13 Australia’s definition of an imported case (Box 1) involves assessment on a case-by-case basis using epidemiological and virological evidence and may be more sensitive and less specific than Canadian and United States criteria. For example, an Australian case associated with travel to the Philippines 23 days before symptom onset was considered imported.27 Despite this less restrictive means of assessment, overestimations are unlikely in Australia because health authorities usually conduct extensive investigations for each case.27 The use of a consistent definition for importation status by health authorities both in Australia and internationally would provide more comparable R estimates. Because the recording of importation status requires minimal additional resources and any underestimation of the proportion imported would only overestimate R, we recommend that countries with low or middle incomes initially focus on collecting data on this variable to monitor their measles elimination status.

R estimates obtained using the distribution of generations of spread were consistently found to provide lower estimates in Australia and in the United States (R: 0.44; 95% CI: 0.36–0.52) than those obtained with other methods.11 One explanation for the lower estimates may be that they have lower validity because the generations of transmission for both countries were derived from an algorithm.11 For example, during the period from 2009 to 2011, the longest outbreak duration in Australia (67 days) was calculated to involve seven generations. However, based on an incubation period of 18 days, the generations of transmission of this outbreak could have been as few as four. Recording the number of generations of transmission would improve the validity of R estimates calculated using these data.

In our study, R estimates from methods 2 and 3 might also have overestimated the true value of R. Nineteen cases with a missing outbreak reference number were assigned to a particular outbreak on the basis of a flowchart (Fig. 1). We also used the flowchart to evaluate whether sporadic cases with an outbreak reference number might have been part of an outbreak. Although two sporadic cases were subsequently linked to an outbreak, the linkage might have been mistaken and may have resulted in an overestimated R. However, because measles is such a rare disease in Australia, it is likely that cases with temporal and geographical similarities are linked. Moreover, the high sensitivity of the surveillance system made the detection of most measles cases likely. Therefore, it is improbable that a large number of cases were missed and that this resulted in an underestimation of R.

Completeness of surveillance data is imperative for obtaining accurate estimates of R. Comparison of our results with those for Canada and the United States reveals that our data were more recent (2009–2011) than those from Canada (1998–2001)15 and from the United States (1997–1999).27 It is likely that the completeness of data for both countries has improved since these studies were conducted. However, to our knowledge these are the most recently published studies. In Australia, over the three-year period analysed, completeness of data on place of acquisition was 100%. Data on outbreak reference number, however, were incomplete, in part because most jurisdictional health authorities do not assign unique outbreak reference numbers to known sporadic cases (Nicolee Martin, personal communication, 2012). If all known sporadic cases were assigned outbreak reference numbers, they could be more accurately differentiated from cases in which a link to an identified outbreak was not determined, thereby providing a more accurate R estimate. However, recording the outbreak reference number and determining the number of generations of spread requires more extensive case follow-up to enable linkage between cases and may be more challenging in countries with low or middle incomes.

The methods used in our study have several potential limitations. We assumed homogeneity in the susceptibility to measles and the mixing of populations.11 Past outbreaks have identified susceptible subpopulations in Australia, including young adults and infants too young to be vaccinated.28–30 Therefore, R may be > 1 in certain subpopulations. However, our results are similar to R estimates using serosurveillance data,26 which take into account population dynamics. Despite these limitations, these methods provide a suitable means to monitor the status of elimination in Australia as a whole, on a regular basis, through the use of high-quality national notification data and could potentially be applied in lower-resource settings too.

Notification data have the potential to play an important role in verifying measles elimination, especially when all WHO regions are working towards eliminating measles by 2020.24–25 Because high-quality surveillance is essential for verification,25 it should be feasible
Estimation of $R$ to track measles elimination in Australia

May Chiew et al.

Research

Finding:

Heather F Gidding is funded by a National Health Medical Research Centre postdoctoral research scholarship. The National Centre for Immunisation Research and Surveillance of Vaccine Preventable Diseases is funded by the Australian Government Department of Health, Canberra.

Competing interests: None declared.

Estimation of $R$ to track measles elimination in Australia

Objectif: Estimer le taux de reproduction efficace de la rougeole ($R$) en Australie en modélisant les données de notification de routine recueillies.

Méthodes: $R$ a été estimé pour la période 2009-2011 à l’aide de 3 méthodes, en utilisant les données du Système de Surveillance des Maladies à Déclaration Obligatoire de l’Australie. La méthode 1 a estimé que $R$ est égal à 1 – $P$, où $P$ est égal à la proportion de cas importés, tels que déterminés à partir des données obtenues sur le lieu de l’acquisition. Les autres méthodes ont estimé $R$ en ajustant un modèle de régression logistique.

Conclusion: There are several ways to estimate the reproduction number ($R$) in settings where routine data are used. The present study shows that all three methods produced consistent estimates of $R$ over the study period in Australia. This can be used to monitor the effectiveness of measles control and elimination efforts in the country.

Résumé

Estimation du taux de reproduction efficace de la rougeole en Australie à partir de données de notification de routine

Objectif: Estimer le taux de reproduction efficace de la rougeole ($R$) en Australie en modélisant les données de notification de routine recueillies.

Méthodes: $R$ a été estimé pour la période 2009-2011 à l’aide de 3 méthodes, en utilisant les données du Système de Surveillance des Maladies à Déclaration Obligatoire de l’Australie. La méthode 1 a estimé que $R$ est égal à 1 – $P$, où $P$ est égal à la proportion de cas importés, tels que déterminés à partir des données obtenues sur le lieu de l’acquisition. Les autres méthodes ont estimé $R$ en ajustant un modèle de régression logistique.

Conclusion: There are several ways to estimate the reproduction number ($R$) in settings where routine data are used. The present study shows that all three methods produced consistent estimates of $R$ over the study period in Australia. This can be used to monitor the effectiveness of measles control and elimination efforts in the country.
processus de ramification sous-critique qui modélise la propagation d’une infection avec un R donné pour les distributions observées de l’amplitude de l’épidémie (méthode 2) et des générations de la propagation (méthode 3). Le logiciel Stata version 12 a été utilisé pour la méthode 2, et le logiciel Matlab version R2012 pour la méthode 3. Pour toutes les méthodes, le calcul des intervalles de confiance à 95% (IC) a été effectué en utilisant une approximation normale basée sur les erreurs type estimées.

Résultats Au cours de la période 2009-2011, 367 cas de rougeole à déclaration obligatoire ont été enregistrés en Australie (taux moyen annuel: 5,5 cas par million d’habitants). Les données étaient complètes à 100% pour le statut de l’importation, mais complètes à seulement 77% pour le numéro de référence de l’épidémie. R a été estimé comme étant < 1 pour toutes les années et tous les types de données, avec des valeurs de 0,65 (IC à 95%: 0,60–0,70) obtenues par la méthode 1, avec des valeurs de 0,64 (IC à 95%: 0,56–0,72) par la méthode 2 et avec des valeurs de 0,47 (IC à 95%: 0,38–0,57) par la méthode 3. Conclusion Le fait que des estimations cohérentes de R ont été obtenues à partir de ces 3 méthodes renforce la confiance dans la validité de ces méthodes pour la détermination de R.

Резюме
Оценка значения эффективного воспроизведения кори в Австралии по данным уведомлений, получаемых стандартным образом

Цель Оценить значение эффективного воспроизведения кори (R) в Австралии путем моделирования получаемых стандартным образом данных уведомлений.

Методы Значение R оценивалось для периода с 2009 по 2011 г. с применением трех методов, предполагающих использование данных австралийской Национальной системы наблюдения за заболеваниями, подлежащих регистрации (National Notifiable Disease Surveillance System). Метод 1 оценивал значение R как 1–P, где P равно соотношению случаев заражения, согласно данным на местах, где были зарегистрированы вспышки. Другие методы оценивали значение R в виде пропорции докритических разветвляющихся процессов, моделировавших распространение инфекции с заданным значением R, в соответствии с наблюдаемым распределением размеров вспышек (метод 2) и формированием разброса (метод 3). Версия Stata 12 использовалась для метода 2; версия Matlab R2012 использовалась для метода 3. Для всех методов вычисление доверительных интервалов (ДИ) в 95 % производилось с использованием нормального приближения на базе оценок средних квадратических ошибок.

Результаты За период с 2009 по 2011 г. в Австралии было зафиксировано 367 случаев заболеваний корью, подлежащих регистрации (среднегодовой уровень: 5,5 случаев на миллион человек населения). Данные были полными на 100% для случаев заражения, но на 77% полными для контрольного значения вспышки заболевания. Значение R было оценено как < 1 для всех лет и типов данных и составило 0,65 (ДИ 95 %: 0,60–0,70) по результатам, полученным с использованием метода 1; 0,64 (ДИ 95 %: 0,56–0,72) по результатам, полученным с использованием метода 2; и 0,47 (ДИ 95 %: 0,38–0,57) по результатам, полученным с использованием метода 3.

Вывод Тот факт, что непротиворечивые оценки значения R были получены по результатам использования всех трех методов, подкрепляет уверенность в обоснованности использования этих методов для определения значения R.

Resumen
Estimar el número de reproducción eficaz contra el sarampión en Australia a partir de los datos de notificaciones rutinarias

Objetivo Estimar el número de reproducción efectiva del sarampión (R) en Australia mediante el modelado de datos de notificaciones recopilados de forma rutinaria.

Métodos Se estimó R para 2009–2011 a través de tres métodos y con los datos del Sistema nacional de vigilancia epidemiológica de enfermedades de notificación obligatoria de Australia. El método 1 estimó R como 1–P, donde P es igual a la proporción de casos que fueron importados, según lo determinado a partir de datos del lugar de adquisición. El resto de métodos estimó R mediante la adaptación de un proceso de ramificación subcrítico que modelaba la propagación de una infección con un R dado a las distribuciones observadas de tamaños del brote (método 2) y a las generaciones de propagación (método 3). Se empleó la versión 12 de Stata para el método 2 y la versión R2012 de Matlab para el método 3. Se realizó el cálculo de intervalos de confianza del 95 % (IC) para todos los métodos por medio de una aproximación normal basada en errores estándar estimados.

Resultados Durante 2009–2011, se dieron 367 casos de sarampión de declaración obligatoria en Australia (tasa anual media: 5,5 casos por millón de habitantes). Los datos estaban completos al 100% para el estado de importación, pero al 77% para el número de referencia del brote. R se estimó como < 1 para todos los años y tipos de datos, con valores de 0,65 (IC del 95 %: 0,60–0,70) a través del método 1, 0,64 (IC del 95 %: 0,56–0,72) con el método 2 y 0,47 (IC del 95 %: 0,38–0,57) con el método 3.

Conclusión La obtención de estimaciones consistentes de R mediante los tres métodos aumenta la confianza en la validez de estos métodos para determinar R.

References
Estimation of R to track measles elimination in Australia

May Chiew et al.


