RAPID COMMUNICATIONS

Relapsing vivax malaria cluster in Eritrean refugees, Israel, June 2010
by E Kopel, E Schwartz, Z Amitai, I Volovik

Spotlight on measles 2010: Measles outbreak in a mainly unvaccinated community in Essen, Germany, March – June 2010
by H Roggendorf, A Mankertz, R Kundt, M Roggendorf

Cluster of Legionnaires’ disease associated with a public whirlpool spa, France, April – May 2010
by C Campese, D Roche, C Clément, F Fierobe, S Jarraud, P de Waelle, H Perrin, D Che

RESEARCH ARTICLES

Type and quantity of data needed for an early estimate of transmissibility when an infectious disease emerges
by NG Becker, D Wang, M Clements

NEWS

Injecting drug use in Europe: stable or declining
by L Wiessing, D Klemppová, D Hedrich, L Montanari, VA Gyarmathy
We report on a cluster of relapsing vivax malaria among Eritrean refugees residing in Israel. Since the beginning of 2010, 15 cases have been identified. Five of the six patients who had complete medical and epidemiological histories, reported Sudan as the place of primary infection during their journey to Israel, and having had the first relapse in Israel, six months later (median). Suggested place of exposure is the region of the Eritrean refugee camps in Sudan.

**Introduction**

Malaria, once endemic in Israel, was eradicated almost 50 years ago, although its vectors, several malaria-transmitting species of *Anopheles* mosquitoes, still exist in various parts of the country [1,2]. Every year between 60 and 100 imported cases of malaria are reported to the Ministry of Health. Most of these cases are travellers returning from endemic countries to Israel and only few of them are immigrants from Sub-Saharan Africa [3].

Currently, there is a cluster of relapsing *Plasmodium vivax* malaria among Eritrean refugees in Israel. The epidemiological investigation, which is summarised here, was conducted by the local health office in Tel Aviv and is therefore limited to the Tel Aviv district.

**Methods**

On 7 June 2010, a cluster of five malaria cases was reported to the Tel Aviv District Health Office. These cases were Eritrean refugees under treatment in the same hospital during the first week of June. An epidemiological investigation was initiated. The case definition was laboratory-confirmed malaria excluding returning travellers.

The following investigation measures were taken.

- Species were identified by thick and thin smears with confirmation by real-time reverse-transcriptase PCR [4].
- Medical records of cases were obtained from hospitals.
- Oral interviews were conducted with four cases by a native speaker of Tigrinya (one of the two working languages in Eritrea). The interviews were based on a short epidemiological questionnaire that contained questions on demographics, the route of the journey to Israel, current and past malaria illness history and possible exposures (time, place, mosquito bites) in Israel or abroad.

- The local health office provided records on past epidemiological data regarding non-traveller cases of malaria in the district since 2006.
- An alert was issued to the hospitals and laboratories in the district, which were requested to report all cases of diagnosed malaria and to confirm them at the national reference laboratory.
- With the assistance of an expert malaria advisor, concise clinical guidelines for proper management [5,6] of cases were promptly issued to local infectious diseases units at all hospitals in the district.

**Results**

Five cases of non-traveller malaria were reported in the district from 2006 to 2009 (Figure). All of these cases had been imported, three of whom were refugees from Eritrea. Since the beginning of 2010, 15 cases, all of them Eritrean refugees, have been reported. Most of them (nine of 15) were diagnosed in the first two weeks of June 2010. No other cases of imported malaria in other migrants were documented during this period of 2010 in the district.

Twelve cases were male (80%) and the median age was 25 years (interquartile range (IQR): 21.5–29 years). All the cases had laboratory-confirmed *P. vivax* infection. Low parasitaemia, ranging from <0.1% to 1% at the time of diagnosis, was demonstrated for 12 patients for whom these data were known.

All 15 patients presented with similar clinical characteristics of intermittent fever which was usually accompanied by shivering and headaches. Eleven patients had mild to moderate anaemia, usually normocytic, and four of 15 patients had splenomegaly.

Thirteen of the 15 cases for whom the onset of illness and date of arrival to Israel were known had arrived in Israel during the six months before that date.
Eritrean refugees have gradually become the predominant group of asylum seekers who enter Israel from the Sinai desert (Egypt). In the first four months of 2010, 3,793 Eritrean refugees entered the country and constitute 82% of all asylum seekers who entered in this period [8].

These figures represent a 40% increase in the Eritrean refugee population in Israel during the first four months of 2010 alone. In addition, a more than fourfold increase in the crude incidence rate of non-traveler malaria in Israel was observed in the same period: in 2009, the crude incidence rate was 0.77 cases per 1,000 migrants, while in the first half of 2010, it was 3.58 cases per 1,000 migrants*. These estimated calculations are based on data from the National Department of Epidemiology in the Israeli Ministry of Health and from an analysis report of the Knesset Research and Information Centre [8]. Consequently, an overall increase in malaria activity along the refugee route in Africa must have played a significant role in this cluster.

Therefore, it is reasonable to assume that a place on the journey is a recent common place of exposure. This place was most probably in Sudan rather than Eritrea as the lowland of Eritrea is mostly affected by P. falciparum and not by P. vivax [9]. Furthermore, a dramatic decline in the incidence rate of malaria was observed in the recent years in Eritrea, due to successful eradication programmes [10]. In fact, almost all cases who had complete medical and epidemiological history, namely a third of the total number of cases, specifically recalled having had the first malaria attack during their stay in Sudan, which was usually a period of two months before arriving to Israel.

Moreover, this cluster may actually reflect the expanding prevalence of a previously reported [11] small focus of vivax malaria in Sudan: 10 of 83 blood samples were PCR-positive for P. vivax (on the background of predominant P. falciparum in this area) in some of the Eritrean refugee camps in eastern Sudan, a frequently flooded low plain region in which malaria may have remained the leading cause of morbidity and mortality.

**Conclusions**

The evident ongoing rise of human reservoirs of malaria in the region may potentiate the risk for the re-emergence of locally acquired mosquito-transmitted malaria in Israel and neighbouring countries. This warrants tight national surveillance for new cases, proper clinical management and follow-up of current cases, and effective control measures of the local *Anopheles* vectors. In addition, it highlights the need for increased malaria surveillance in the refugee camps of eastern Sudan.
Acknowledgements

We thank Anat Scheffer, Sofia Katser and Sara Germay for their assistance in the epidemiological investigation. We also thank the epidemiology unit of Tel Aviv Sourasky Medical Centre for the initial notification and the National Reference Centre for Parasitology for the cooperation.

*Authors' correction

The following correction was made on 5 July 2010 on request of the authors: in the Discussion section, the second sentence of the second paragraph: ‘in 2009, the crude incidence rate was 0.77 per 1,000 cases, while in the first half of 2010, it was 3.58 per 1,000 cases.’ was replaced with the following sentence: ‘in 2009, the crude incidence rate was 0.77 cases per 1,000 migrants, while in the first half of 2010, it was 3.58 cases per 1,000 migrants.’

References

On 15 March 2010, a case of measles was reported to the District Health Office in Essen. In total 71 cases occurred from 15 March to 19 May (four cases hospitalised), with the majority linked to a Waldorf school. Only one case had been vaccinated twice, two cases had been vaccinated once. Immediate and consequent exclusion of non-immune children from classes for two weeks as well as the adjacent spring break prevented the wider spread of the virus.

Introduction
Measles outbreaks occurred in Germany in recent years [1,2] despite the recommendation of the German Standing Committee on Vaccination (STIKO) to vaccinate all children with two doses of the measles-mumps-rubella (MMR) vaccine, according to the vaccination schedule. Vaccinations are not mandatory in any of the German laender but vaccination cards are routinely checked at a medical examination at school entry.

The last major measles outbreak in Germany in 2006 [2] involved 2,300 cases. During this outbreak, 414 (18%) of the infected children in the town of Duisburg (North Rhine Westphalia) were hospitalised and two died with severe encephalitis. Since then, information and vaccination programmes were enforced in Germany and the number of reported cases declined to 574 in 2009 [3]. In 2008, countrywide vaccination coverage for measles (with two doses) in six-year-olds was 89% [4] whereas in Essen, vaccination coverage was 92% in the age group 11–13 years (unpublished data).

Outbreak description
On 15 March 2010 one serologically confirmed measles infection in a 13-year-old student and four clinical suspected cases from a Waldorf school in Essen were reported to the District Health Office.

By 19 May 2010, a total number of 71 cases (68 children and three adults) were reported to the District Health Office. Up to nine cases were reported on a single day but after the spring break from 29 March to 9 April only a maximum of two cases were reported per day. Twenty-eight pupils (39%) infected in this outbreak are aged from 11 to 15 years, followed by 19 pupils (27%) aged between 0 and 5 years, four of whom were not eligible for vaccination. Eighteen children (25%) are aged between 6 and 10 years and two students between 16 and 20 years. All cases had epidemiological links apart from the three adults (4%), who were aged over 20 years. For one case the age was not known.

Sixteen of 71 reported cases were serologically confirmed. Genotyping of two isolates by the National Reference Centre revealed Genotype D8. A number of parents refused to have their children tested and this limited the outcome of our investigation.

Three children and one adult were hospitalised with fever and severe rash but they did not develop any major complications.

According to the information received from parents of the 71 cases, 30 could be identified as members of the Waldorf school or kindergarten, 18 as siblings of these members and 20 as visiting doctors who do not recommend vaccination. However, the members of the Waldorf school and their siblings might also visit doctors who do not recommend vaccination. The three adults did not have any link to any of these groups.

Seven siblings have not been seen by a doctor and therefore, have not been reported but they were included in the analysis with the date of onset of symptoms. Another seven cases attended different schools or kindergartens or were too young to attend one. Therefore 25 secondary cases occurred in other schools and kindergartens in unvaccinated children, whose parents refused vaccination. The number of infections through household contacts was 30. One case contracted the disease after vaccination with one dose and another one after vaccination during the incubation period. One of the adults (aged 28 years) had received two vaccinations against measles at the age of two years. All the other 68 cases had not been
vaccinated at all. Only three cases could not be allocated to one of the groups (Waldorf school, sibling or attending doctor who did not recommend vaccination) indicating that this outbreak was mainly restricted to the above-described groups.

In addition to the 71 cases reported in Essen, five cases reported in neighbouring cities could be traced back to contacts with children from the Waldorf school in Essen. Furthermore, one case reported to the District Health Office in Sonthofen (southern Germany) is linked to a child from Essen who spent his holidays there. Another case from Zwickau (eastern Germany) could have been exposed while visiting a paediatric practice in Essen that did not recommend vaccination but was seeing measles patients at this period.

**Public health intervention**

In order to stop this outbreak and to protect the non-immune children, and since this outbreak involves a school with low vaccination coverage against measles, measures to prevent the spread of the infection according to the national Protection Against Infection Act (Infektionsschutzgesetz) were immediately enforced [6]. This included obligatory control of vaccination certificates and exclusion of non-immune students from classes for 14 days. A firm recommendation for vaccination with a first or second dose against measles was given.

The measures started on 15 March 2010, when the school administration was advised by the District Health Office to hand out leaflets in order to inform the parents about the measles outbreak and the measures planned and recommending that children stay at home if they develop symptoms. Parents were asked to have their children vaccinated if they had not received two doses of MMR; in case of non-compliance, the children were excluded from classes for two weeks. On the following day, staff from the District Health Office (two paediatricians, two health supervisors and two assistants) checked the vaccination certificates of all the pupils attending school on that day, before the beginning of lessons. The control of the vaccination certificates showed that 311 of 762 children (41%) attending the Waldorf school were not vaccinated against the disease or had not had measles before. None of the susceptible students attended classes. However, of the 311 non-vaccinated pupils, 30 (10%) contracted the disease in the following four weeks. Some children had already contracted the disease before the index case but had not been reported earlier. The investigations revealed that the first patient had shown symptoms on 3 March and another six cases followed until the first serologically confirmed case was reported.

Only children who had been immunised against the disease or who had a history of previous disease were allowed to attend lessons. All the others were sent home and the parents were recommended to have their children vaccinated. Following this recommendation, four children were vaccinated. Information on teachers’ and other school staff’s immunisation status was also available and it was communicated to the District Health Office: all were immune.

The school administration and the teachers were very cooperative in the organisation of the vaccination certificates control. However, the majority of parents indicated clearly that they disagreed with having their children vaccinated against measles.

All the paediatricians in the area were informed by email about the outbreak. The population of Essen was informed via newspapers, Internet and local television. Staff from the District Health Office had several discussions as well as conversations via email with parents who were concerned about the exclusion of their children from school. In the end, the necessity of these measures was agreed upon at a meeting with parents’ representatives and staff from the District Health Office held on 25 March 2010.

**Discussion and conclusions**

Of the 71 cases in this outbreak, only one had received two doses of MMR and a further two cases had received only one dose. Given the high rate of second-dose MMR vaccination coverage (92%) in six-year-old pupils during the school entry examination of 2009, as well as in 12-year-olds by annual control of vaccination certificates [7], we hope that the outbreak will stop soon and not extend far beyond the Essen Waldorf community, which has a critical attitude towards vaccination. However, seven new cases were reported in late June in Essen, who have no epidemiological link to the outbreak in the Waldorf community. Immediate temporary exclusion of children without measles vaccination or naturally acquired immunity from classes has helped to prevent the spread of the virus to a larger number of children. The cases that occurred during spring break had had the incubation period before the break, and the spring break might have contributed to the decrease in the number of newly reported cases.

The virus detected in this outbreak is very similar to a virus imported from India, which caused an outbreak at a Waldorf school in Berlin at the beginning of 2010 [8], but it is not identical (one sequence variation). Therefore, a link to the current outbreak is possible but could not be confirmed.

The goals of the World Health Organization (WHO) to eliminate measles [9,10] cannot be achieved as long as doctors do not recommend vaccination or parents refuse to have their children vaccinated against measles. More efforts and a 95% coverage with two doses of MMR vaccine in children are needed for measles eradication in Germany, in order to meet the WHO goals [11].
**Acknowledgements**

We thank all the medical staff and Health officers from the Department of Infectious Diseases Prevention, District Health Office, Essen, for their support.

**References**


In May 2010, a cluster of three cases of Legionnaires’ disease was identified in France. The results of the epidemiological, environmental and microbiological investigations allowed the rapid identification of a public whirlpool spa as the most probable source of contamination and the implementation of appropriate control measures. This investigation has stressed the need for good cooperation between partners and the importance of the molecular analysis of Legionella strains.

Background
Legionnaires’ disease is an atypical pneumonia caused by the inhalation of aerosols contaminated by Legionella [1]. Legionella are ubiquitous bacteria and can grow in natural and man-made environments. Aerosol-generating devices such as wet cooling towers and water systems are well documented as sources of Legionnaires’ disease [2]. Spas have also been widely acknowledged as a source of exposure in outbreaks [3-5].

In May 2010, three cases of Legionnaires’ disease were registered by the local health authority in a district in the north-east of France (the Ardennes). These cases had visited the same spa centre during the 10-day period before the onset of symptoms. This paper describes the cluster, the investigation and the control and prevention measures implemented.

Methods
In France, notification of Legionnaires’ disease is mandatory and the local health authority is in charge of the implementation of epidemiological and environmental investigations. Cases or their relatives are systematically interviewed using a standardised questionnaire. The objectives are to assess risk factors for contracting the disease, to identify a possible source of exposure and to rapidly detect clustered cases or outbreaks, in order to implement appropriate prevention and control measures. If necessary, the environmental team of the local health authority investigates the potential sources of contamination and collects water samples for laboratory analysis. Legionella strains from clinical and environmental samples are analysed by the National Reference Centre for Legionella. Three methods are used: international monoclonal antibody sub-grouping [6], pulsed-field gel electrophoresis (PFGE) typing according to standard procedures [7] and sequence-based typing using standard European procedures [8].

Results
Descriptive epidemiology
The three cases were confirmed cases according to the national [9] and European Union case definitions [10]. All patients had symptoms of acute lower respiratory tract infection and were hospitalized.

Case 1
In early May 2010, the local health authority of a district in the north-east of France (the Ardennes) received notification of this case, a woman in her early 70s, presenting with an underlying chronic disease that increases risk of Legionnaires’ disease. She was diagnosed only by a positive urinary test (no strain was isolated from a respiratory specimen). This patient had visited a spa centre towards the end of April, without using the whirlpool spa; she used a sauna located inside the room containing the whirlpool spa. She developed symptoms three days after visiting the spa centre and recovered after treatment with antibiotics.

Case 2
Four days after the notification of Case 1, a second case was notified: a woman in her early 50s. Immediate interview was not possible. She was diagnosed by a positive urinary test and presented tobacco smoking as a risk factor. One respiratory specimen was collected and a Legionella strain was isolated. Two days after notification, she died, despite intensive treatment including antibiotics. Her relatives were interviewed the day after she died and indicated she had visited the same spa centre six days after Case 1 had and had stayed in the room containing the whirlpool spa.
Case 3
Sixteen days after the notification of Case 1, a third case was notified: a man in his early 30s, presenting no risk factors. He was diagnosed by positive urinary test. A respiratory specimen was also available for this case and a *Legionella* strain was isolated. The patient had visited the same spa centre as Cases 1 and 2 had and had used the whirlpool spa. He had visited the spa centre 17 days after Case 1 had and developed symptoms (digestive disorders first, and then cough and fever) four days after his visit. The case recovered after treatment with antibiotics.

Investigation and control measures
Three days after the first case notification, an environmental investigation was carried out at the spa centre. Water samples from the whirlpool pump output were collected and control and prevention measures were advised. Three days after the second case notification, the local heath authority decided to stop the use of the whirlpool spa. The national and regional health authorities recommended closing the spa centre two days later. In parallel, as the three cases lived in the same area, an investigation was undertaken to look for other possible sources of contamination. Three cooling towers were investigated: one was closed during the exposure period of the three cases; in the other two cooling towers, routine control samples were negative for *Legionella*.

A decision to undertake active case finding was taken following the notification of the second case. A request for immediate notification of new cases was sent to the local hospitals and general practitioners in the vicinity of the spa. Five days after notification of the second case, the local administrative authority issued a press release in order to inform people who might have visited the spa centre during the previous 14 days and to encourage them to visit their general practitioner if they developed symptoms. Case 3 visited his general practitioner after receiving this information. No other case has been notified.

Analysis of the samples from the whirlpool spa, available the day after use of the whirlpool spa was stopped, showed contamination with *Legionella pneumophila* serogroup 1 (Lp1) at a level of 150,000 colony-forming units per litre. Five *Legionella* strains from environmental samples were sent to the National Reference Centre for *Legionella* for genomic analysis and comparison with the two strains from the clinical samples (Cases 2 and 3). The Lp1 strains from the clinical and environmental samples shared the same characteristics: they were indistinguishable by monoclonal antibody subgrouping (Allentown/France) and by sequence-based typing (sequence type 23) and they had the same PFGE profile. These strains were not considered as endemic strains: among more than 2,500 clinical isolates collected during the last 10 years and typed by the National Reference Centre, only one, isolated in 2009, had the same PFGE profile associated with monoclonal antibody subgroup Allentown/France and sequence type 23. The corresponding patient had not visited the spa implicated in the April – May 2010 cluster.

Discussion and conclusions
The epidemiological, environmental and microbiological investigations allowed the rapid identification of the whirlpool spa as the most probable source of this cluster and the implementation of control measures. On the day the relatives of Case 2 were interviewed, the local health authority immediately decided to stop the use of the whirlpool spa.

The speed of a response may influence the dimension of clusters and outbreaks. Several conditions must be satisfied in order to rapidly identify and control a source of contamination. Firstly, cases have to be notified as soon as the diagnosis is established and then interviewed as soon as possible, in order to be able to build a hypothesis about the potential source(s). For this purpose, it is essential to have a standardised questionnaire, which helps to gather all the relevant information about the exposure. Secondly, it is essential to have clinical specimens for comparing the genomic profiles of the strains and to obtain strong evidence for confirming the source of contamination. In France, bacterial culture of respiratory samples is recommended for all cases with a positive urinary test and clinicians are regularly reminded about this recommendation.

The percentage of cases for which a *Legionella* strain has been isolated is stable in France, at around 18% during the past 10 years. Even if the culture result often does not change the diagnosis and treatment – particularly for cases with Lp1 infection diagnosed by urinary test – there should be further effort to increase the number of strains isolated from cases. Otherwise, as a result of the delay between exposure and the recognition of a cluster, the conditions in the suspected sources may have changed by the time they are investigated, thus limiting the capacity to identify the source of contamination [11]. For example, whirlpool spas may have been repeatedly drained and disinfected and analysis of water samples could remain negative [12]. The local team responsible for environmental investigation must be informed and trained in the best procedures for sampling such installations and equipment (e.g. swab samples of biofilms and samples of aerosol collections).

This cluster highlights once again the importance of ongoing vigilance regarding the proper maintenance of the water in spa pool facilities and the importance of a reactive surveillance system.

References

www.eurosurveillance.org


An early estimate of disease transmissibility is essential for a well-informed public health response to a newly emerged infectious disease. In this study, we ask what type and quantity of data are needed for useful estimation of the initial reproduction number (R). It is possible to estimate R from case incidence data alone when the growing incidence of cases displays a wave pattern, because the pattern provides information about the serial interval (the time elapsed between the onset of symptoms of a case and symptom onset in individuals infected by that case). When the mode of the serial interval distribution is small, 1.5 days or less, there is generally no informative wave pattern in the observed series of daily incidences. The precision of the estimate of R is then improved substantially by having some observations on the serial interval. For an infectious disease with characteristics such as those of influenza, an estimate of R able to inform plans to mitigate transmission is obtained when the cumulative incidence of cases reaches about 300 and about 10 observations on the serial interval are available.

Introduction

Concern about the risk posed to humans by avian influenza A(H5N1) encouraged substantial planning for the possible emergence of pandemic influenza [1-4]. The emergence of the pandemic influenza A(H1N1) strain in 2009 further highlighted the importance of pandemic preparedness. Key elements of preparedness plans are disease transmissibility, the rate of disease progression and how these change with use of antiviral drugs, vaccines and social-distancing measures. Assumptions about disease transmissibility and progression are necessarily based on data from past pandemics and seasonally circulating influenza strains, but a future pandemic strain may have quite different characteristics.

Preparedness planning deals with this uncertainty by assessing the effectiveness of interventions under different scenarios. When a new viral strain emerges, it is important to determine which scenario obtains, because the effectiveness of some interventions is scenario dependent. For example, targeted use of antiviral drugs, early and liberally, may contain an influenza strain with a modest transmission rate, but would be ineffective against a highly transmissible strain. Timeliness is also important, because an intervention is most effective when introduced early. Here we consider what data from the early stage of an outbreak, and how much, are needed to inform decisions about interventions needed to mitigate the impact of a pandemic to a manageable level.

It is convenient to quantify disease transmissibility by R, the effective reproduction number of infective individuals. At any time, R is the mean number of infections generated by a ‘typical’ infective person, given current levels of immunity and public health interventions. It quantifies the growth in the number of cases from one generation of cases to the next. We aim to estimate the initial R from early incidence data of an outbreak. Incidence data alone seem inadequate for this estimation: we also need information about the serial interval, the time elapsed between the onset of symptoms of a case and symptom onset in individuals infected by that case. The artificial incidence series A and B of Table 1 illustrate this point. Comparing incidences on days 0, 2, 4, 6 and 8 suggests the two outbreaks are growing similarly over time, while comparing cumulative incidences suggests series B is the larger threat. However, series A actually poses the greater threat (larger eventual attack rate) because it is consistent with R = 4 and every infected person having a short symptomatic infectious period on the second day after infection, while series B is consistent with R = 2 and a short symptomatic infectious period on the first day after infection. In other words, reproduction numbers corresponding to incidences that appear to be growing similarly can differ by a factor of two when the mean serial interval differs by a factor of two. This shows that estimates of R obtained by assuming a form for the serial interval distribution come with the risk of substantial estimation bias.

The basic reproduction number (R_0) is the mean number of infections generated by a ‘typical’ infective person in a community with everyone susceptible and no public health interventions in place. Throughout this paper, R
refers to the initial reproduction number. For pandemic influenza this is likely to differ from \( R_0 \) for two reasons. Firstly, some cross-immunity from previous exposure to influenza strains may be present. Secondly, prior alertness to the possibility that the pandemic strain may be imported, and its unknown severity, may result in atypical behaviour and an enhanced public health response.

Wallinga and Teunis [5] provide a method for estimating \( R \) that is based on considering, for every case, who might have been responsible for that infection. The distribution of the serial interval is assumed to be known. Cauchemez et al. [6] modified the approach to enable dynamic estimation of \( R \) over time. The above comparison for case series A and B suggests that it is preferable to estimate \( R \) and the mean serial interval simultaneously from early data. A method for making Bayesian inferences about \( R \), without assuming a specific distribution for the serial interval, is proposed by Cauchemez et al. [7]. They assume that a certain fraction of infections are traced as an epidemic progresses. An alternative approach to estimating \( R \) during the early stage of an epidemic is described by White and Pagano [8]. Their results suggest that it is possible to estimate \( R \) and parameters of the serial interval distribution simultaneously using only daily incidence data. Inspection of case series A of Table 1 indeed suggests that there is scope to estimate \( R \) from incidence data alone when the pattern of incidences is strongly suggestive of the serial interval. When growth in incidence exhibits waves over time, we can regard the sources of infection in one wave to be the cases of the previous wave, as illustrated for smallpox by Becker [9]. Here we consider estimation of \( R \) by both maximum likelihood and Bayesian methods. The aim is to determine what data are needed to make the estimate of \( R \) precise enough to inform decisions on public health interventions.

**Methods**

The alert of a possible pandemic virus strain instigates enhanced surveillance of incoming travellers and the general population. It is therefore possible to have daily incidence data of reasonable quality during the early stage of a detected outbreak. Observations on serial intervals are harder to collect because it is often difficult to ascertain the source of an infection. However, the first cases of a newly emerged infection are often travellers and subsequent local cases can sometimes be linked to incoming infected travellers. This can provide observations of serial intervals, as can sequential cases in early household outbreaks.

For maximum likelihood estimation and Bayesian inference of \( R \), we need a likelihood function. We use the likelihood function proposed by White and Pagano [8], which is based on the infection process depicted in Figure 1, with one modification. We augment the likelihood with a contribution for independent observations of the serial interval, as described in the Appendix. We also use an unrestricted range of distributions for the serial interval, so we can better explore how results depend on the shape of this distribution. This likelihood function was used for maximum likelihood estimation and in Bayesian inferences via Markov chain Monte Carlo (MCMC) methods on simulated data [10], to see how these inferences perform with different amounts of data and with different rates of disease transmission and progression.

For our assessment of data needs we simulated, for each choice of parameter values, a large number of outbreaks, as in White and Pagano [8]. Specifically, we begin an outbreak with a fixed number of newly infected individuals. We assume that the number of infections generated by an infected case has a Poisson distribution. This assumes that each case has the same potential to infect others. We also assume that the serial interval has a multinomial distribution, as depicted in Figure 1.

We covered values of \( R \) in the range one to five, and a wide range of plausible shapes for the distribution of the serial interval.

---

**Table 1**

<table>
<thead>
<tr>
<th>Incidence series</th>
<th>Daily incidence counts</th>
<th>Day*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-1 0 1 2 3 4 5 6 7 8</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>– 1 0 4 0 16 0 64 0 256</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>– 1 2 4 8 16 32 64 128 256</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>– 10 0 20 0 40 0 80 0 160</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>4 6 8 12 16 24 32 48 64 96</td>
<td></td>
</tr>
</tbody>
</table>

*Day 0 is the day initial cases present. In series D, four additional initial cases present on the previous day.

Four series of daily incidence counts that coincide with the mean count when \( R = 4 \) for series A, \( R = 2 \) for series B, C and D, and a short symptomatic infectious period on the first day after infection in series B and on the second day after infection in series A, C and D.

**Figure 1**

The infection process: mean number of secondary cases generated by a single case

\( p_i \); probability that a serial interval is i days; \( R \): initial reproduction number; \( t \): day of symptom onset.

Note that \( R_{po} \) is simply \( R \) multiplied by \( p_i \).

Mean number of secondary cases, with onset of symptoms in the next three days, generated by a single case with onset of symptoms on day \( t \).
How precise should the estimate of $R$ be? We note that a precise estimate is valuable when $R$ is near one, because it is then useful to be assured that a small amount of additional intervention, such as use of antiviral drugs or restricted school attendance, may contain the outbreak. When $R$ is large (e.g., $R > 3$), we know that considerable intervention is required and the precise value of $R$ is not quite as critical. On this basis, we aim for a precision so that the lower and upper values of a 95% credibility interval lie 25% or less below and above the value of $R$, respectively.

**Results**

As mentioned, data needs were investigated by assessing estimates obtained from many simulations of randomly generated outbreaks. Illustrative results for such simulated outbreaks are given for a few combinations of parameter values in Table 3 in the Appendix. This comprehensive assessment of the methods of inference from such simulated outbreaks led to several useful findings. Here we report these findings with reference to four simple illustrative incidence series, specifically chosen to point to the underlying reasons for the results.

First, we found circumstances when a useful estimate for $R$ can be obtained from daily incidence data alone and having independent observations on the serial interval does not improve the precision of the estimate appreciably. This point is illustrated by estimating $R$ from the case incidences shown in series A of Table 1. Let $p_i$ denote the probability that a serial interval is $i$ days. Incidences A coincide with the mean incidence counts, when $R = 4$ and the serial interval is two days (i.e., $p_2 = 1$). The mean serial interval ($\mu$) is then two. Without additional observations on the serial interval, Bayesian inferences (described in the Appendix) gave the 95% credibility intervals for $R$, $p_1$, $p_2$, $p_3$ and $\mu$ shown in row one of Table 2. Note that:

- an $R$ value of four lies in the 95% credibility interval for $R$ and the interval bounds are only about 10% below and above four
- a $\mu$ value of two lies in the 95% credibility interval for $\mu$
- the large value for $p_2$ and small values for $p_1$ and $p_3$ are indicated well by the inferences.

Specifically, note the tight credibility interval for $\mu$, although no independent observations on the serial interval are included.

The above illustration is for incidences artificially chosen to coincide with the mean incidence counts, when $R = 4$ and $p_2 = 1$. Similar performance was observed when incidences were simulated to include a chance component (see Table 3 in the Appendix, for an illustration). The conclusion that incidence data alone can provide useful estimates also holds for variable serial intervals. This is illustrated by results in Table II of White and Pagano [8], who assume certain gamma distributions for the serial interval.

Row two of Table 2 shows the credibility intervals obtained when, in addition, there are 20 observations on serial intervals consisting of 18 serial intervals of two days and one serial interval of each of one day and three days. It is seen that adding the independent observations on serial intervals does not improve the precision of inferences. A similar conclusion is reached from the properties of maximum likelihood estimates. Specifically, the large sample standard deviation of the maximum likelihood estimator for $R$, with parameter values as for series A, is the same (to four decimal places) whether the number of observations on the serial interval is zero or 20.

The extreme pattern of incidences in A is very suggestive of a mean serial interval of two. More generally, we found that incidence data alone provide a good estimate whenever the serial interval distribution is unimodal and the mode is greater than one day. With such a serial interval distribution, a wave pattern tends to be superimposed on the exponentially growing incidence counts, and this pattern is informative about the mean serial interval. In particular, the four gamma

<table>
<thead>
<tr>
<th>Row</th>
<th>Incidence series</th>
<th>$R$</th>
<th>$p_1$</th>
<th>$p_2$</th>
<th>$p_3$</th>
<th>$\mu$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>3.64–4.46</td>
<td>0.00–0.01</td>
<td>0.95–1.00</td>
<td>0.00–0.04</td>
<td>2.00–2.04</td>
</tr>
<tr>
<td>2</td>
<td>A + 20^a</td>
<td>3.66–4.52</td>
<td>0.00–0.02</td>
<td>0.94–0.99</td>
<td>0.00–0.05</td>
<td>1.99–2.05</td>
</tr>
<tr>
<td>3</td>
<td>B</td>
<td>2.11–4.88</td>
<td>0.10–0.88</td>
<td>0.01–0.70</td>
<td>0.01–0.71</td>
<td>1.16–2.51</td>
</tr>
<tr>
<td>4</td>
<td>B + 20^a</td>
<td>1.97–2.59</td>
<td>0.67–0.95</td>
<td>0.01–0.23</td>
<td>0.01–0.21</td>
<td>1.07–1.52</td>
</tr>
<tr>
<td>5</td>
<td>C</td>
<td>1.81–2.25</td>
<td>0.00–0.01</td>
<td>0.97–1.00</td>
<td>0.00–0.03</td>
<td>1.99–2.02</td>
</tr>
<tr>
<td>6</td>
<td>D</td>
<td>1.40–2.06</td>
<td>0.27–0.91</td>
<td>0.01–0.55</td>
<td>0.01–0.51</td>
<td>1.12–2.14</td>
</tr>
<tr>
<td>7</td>
<td>D + 10^a</td>
<td>1.56–2.14</td>
<td>0.15–0.56</td>
<td>0.31–0.76</td>
<td>0.02–0.31</td>
<td>1.50–2.08</td>
</tr>
</tbody>
</table>

$\mu$: mean serial interval; $p_i$: probability that a serial interval is $i$ days; $R$: initial reproduction number.

^a The four artificial incidence series (A–D) in Table 1.

^b The number after the plus sign is the number of observations on the serial intervals.
distributions used by White and Pagano [8] are unimodal and have a mode greater than one day, which is what enables the incidence data alone to produce useful estimates.

In contrast, we also found circumstances when observations on daily incidence data alone are inadequate for simultaneous estimation of \( R \) and parameters of the serial interval distribution. We illustrate this observation by estimating \( R \) from the case incidences shown in series B of Table 1.

The daily incidences in series B coincide exactly with the mean incidence counts when \( R = 2 \) and \( p_1 = 1 \), so that \( \mu = 1 \). These parameter values are not recovered well by Bayesian estimation applied to the incidence data alone, as shown by the credibility intervals in row three of Table 2. The interval for \( R \) is wide and does not contain the value \( R = 2 \). Inferences about the distribution of the serial interval do not suggest a value near one for \( p_1 \), nor for \( \mu \). The four gamma distributions assumed for the serial interval by White and Pagano [8] do not reveal this weakness in making inferences from incidence data alone. By adding 20 independent observations on serial intervals (18 serial intervals of one day, one of two days and one of three days), the width of the credibility intervals narrows appreciably (see row four of Table 2). The main reason for the poor inference when there are no observations on serial intervals lies in the fact that the growing incidence in series B displays no wave pattern, so the incidence data provide minimal information about the mean serial interval. More generally, we found that the precision of estimates of \( R \) from incidence data alone is poor when the probability that serial interval is less than two days exceeds 0.5. Specifically, with a gamma distribution for the serial interval (as in [8]), estimation is poor when the mode of the distribution is zero, e.g. the exponential distribution. In such instances, estimation improves substantially by adding observations on the serial interval.

The following is a useful warning about choosing a suitable value for the number of initial infected individuals in simulation studies that assess methods for estimating \( R \). It is natural to avoid very small outbreaks in simulation studies because they provide little information for estimation and in practice are unlikely to lead to attempts to estimate \( R \). It is therefore common practice to start a simulated transmission chain with a larger number of initial cases. For example, White and Pagano [8] and Cauchemez et al. [6] generally start with 10 initial cases, and sometimes with 100. We found that assessing inferences based on 10 cases on the initial day tends to suggest better precision than is likely with more realistic initial case clusters. We illustrate this point by comparing inferences for case series C and D of Table 1. Incidence series C begins with 10 cases on day 0, while incidence series D begins with six cases on day 0 and four cases with onset of symptoms the previous day. With those respective initial cases, incidence series C and D coincide exactly with the mean counts when \( R = 2 \) and \( p_1 = 1 \). Both series have the same number of cases over the 10-day observation period, so the two series might be expected to contain approximately the same amount of information. The credibility intervals shown in row 5 and row 6 in Table 2 show that inferences for incidence series C are more precise than those for series D. Incidences C lead to better precision, particularly for \( p_1, p_2, p_3 \) and \( \mu \), because the 10 initial cases generate a better wave pattern on the exponentially growing incidences than does the initial case cluster in series D.

Given that incidence data alone are insufficient for estimating \( R \) for all plausible incidence series, it is important to determine how many observations on serial intervals are necessary to estimate \( R \), when incidence data of an outbreak are inadequate for such estimation. Analyses based on Bayesian inferences and on maximum likelihood estimation indicate that just a few observations lead to a substantial improvement in the precision of estimates. This is illustrated in Figure 2. The solid curve shows the large-sample standard error of the maximum likelihood estimate of \( R \) as the number of observations on the serial interval increases. For this curve, we started with 10 cases on the first day and observed the incidence over the following four days, assuming \( R = 2 \) and that the serial interval has the distribution given by \( p_1 = 0.61, p_2 = 0.32 \) and \( p_3 = 0.04 \) (a distribution of the binomial form). The dashed curve shows the standard deviation for the posterior distribution of \( R \) when we have two initial infective cases and incidence counts of 2, 4, 7, 12, 21, 36, 61 and 104 over the next eight days. These counts are the mean counts (rounded to the nearest integer) when

**Figure 2**

*Effect of increasing the number of observations on the serial interval*

Large-sample standard deviation of the maximum likelihood estimate (solid line) and standard deviation of the posterior distribution (dashed line) of the initial reproduction number (\( R \)) as the number of observations on the serial interval increases.
It remains to ask how long a series of incidence data needs to be observed before we can estimate R with useful precision. This depends on the value of R and on the distribution of the serial interval. However, useful guidance is found by noting that estimating R corresponds to estimating the mean of the ‘offspring’ distribution, and so the number of infective individuals who are ‘parents’ is key to answering that question. As generations are not identified, some idea about the mean serial interval is needed. We found that R can be estimated with useful precision if we wait until the cumulative incidence reaches 150 and then continue to observe incidence for a number of days equal to the mean serial interval. Then the incidence data will include close to 150 parents (primary cases). This is illustrated by the results in row 5 (series C) in Table 2 when the incidence data are informative about the serial interval and by the results in row 6 (series D+10) when the incidence data contain little information about the serial interval. Note that series C and D each include 150 parents (primary cases) and there are 160 cases in the final generation whose offspring have not yet been observed.

Discussion

For a disease such as severe acute respiratory syndrome (SARS), with a latent period of a few days and onset of symptoms at about the start of the infectious period, it is very likely that the modal value of the serial interval is located a few days past the point of infection. Our results indicate that R can then be estimated quite effectively from daily incidence data alone. In contrast, for influenza the latent period and time to onset of symptoms tends to be quite short and individuals are thought to be infectious prior to onset of symptoms. It is not clear that the serial interval for the next influenza strain will have a modal value greater than one day. It is therefore sensible to include plans for observing some serial intervals into preparedness plans for pandemic influenza. As few as 10 observations can improve the precision of the early estimate of R substantially. The point that serial interval data improve the estimation of R was also made in the recent paper by White et al. [11].

In contrast to the approach of Cauchemez et al. [7], our approach assumes we have independent observations on the serial interval. This assumption made it feasible for us to carry out the analysis for many choices of parameter values. The assumption has no impact on results for the performance of estimates without serial interval data. For results that include serial interval data, we note that the independence assumption holds when the serial interval data come from a different location. When the serial interval observations are part of the locally collected incidence data, there is some dependence that is ignored by our approach. This is unlikely to have a significant impact on results, since we are assuming we have serial interval observations for less than 10% of infections.

The difficulty of observing serial intervals is exacerbated by the fact that the serial intervals actually observed may not be truly representative of randomly selected serial intervals, because they often arise from household contacts (with higher rates of contact within households) and from infected travellers (who may not have spent all of their infectious period locally).

It is important to be aware that biases may arise from the use of early incidence counts. First, it is important to allow for imported infections. Each imported case must be considered to have been infected elsewhere and not an offspring of a case from an earlier day. The methods used here are easily adapted to allow for this. Second, if a newly emerged infection is not detected immediately there may be a build-up of cases who are then detected in quick succession. Such a burst in the number of detected cases may not reflect the natural history of the infection and disease progression and can lead to initial estimates being biased.

As mentioned previously, the inferences reported here assume that the number of infections generated by an infected case has a Poisson distribution. This assumes that each case has the same potential to infect others and does not allow for variation in infectivity between individuals.

Our assessment clearly involved assumptions. During the early stages of the next newly emerged pandemic strain, it will not be known how appropriate these assumptions are. It will nevertheless be very useful to use these results on the type and quantity of data needed for guidance in preparedness plans for future emerging infections.

Appendix:

http://nceph.anu.edu.au/Staff_Students/Staff%20Publications/Appendix_Becker.pdf

Acknowledgements

This work was supported by Australian National Health and Medical Research Council (NHMRC) grants 471436 and 585536.
References

cenow/2007/04/25-03.html
combinomics.com/News/04180803/H5N1_Pre_Pandemic_Japan.html
The European Monitoring Centre for Drugs and Drug Addiction (EMCDDA) estimates that there may be currently between 750,000 and 1,000,000 active injecting drug users (IDUs) in the European Union (EU). Indirect indicators such as data from drug treatment entrants suggest that the trend in injecting is stable or declining. Data from studies among IDUs, however, suggest that recent recruitment into drug injecting continues in several countries. Injecting drug use is strongly associated with severe health problems in drug users, including both blood-borne infections (e.g., HIV/AIDS, viral hepatitis) and overdose. These results are presented in the report on trends in injecting drug use in Europe recently published by the (EMCCDA) [1].

According to this report, injecting drug use in Europe is mostly linked to opioid use, but now less than half (45%) of those entering treatment for primary opioid use in the EU report ‘usually injecting’ their drug. Between 2002 and 2007, 10 of 26 countries showed a decline in the proportion of injectors among heroin users entering treatment for the first time, likely reflecting a decline in injecting in the untreated population of heroin users. Only two countries, Bulgaria and Slovakia, showed a reverse trend with statistically significant increases. Over the same time period, a declining trend in injecting was observed among primary cocaine users entering drug treatment for the first time, whereas this trend was stable among the primary amphetamine users.

Available estimates of IDU prevalence range from less than one to 15 IDUs per 1,000 population aged between 15 and 64 years, suggesting considerable differences in prevalence between countries. For the 12 EU Member States with prevalence estimates, the weighted average is about 2.5 cases per 1,000 population aged between 15 and 64 years. This figure, if extrapolated to the whole EU, would correspond to between threequarters of a million and one million active IDUs. Trends in IDU prevalence estimates for the period from 2002 to 2007 are only available for five countries and mostly suggest stable prevalence.

The proportion of new drug users (drug use for less than two years) and young drug users (under 25 years) within samples of IDUs recruited in the context of infectious disease surveillance may provide an indirect indicator of recent initiation to injecting. New injectors make up less than 10% of injectors sampled in 10 EU Member States. However, in two EU Member States (Czech Republic and Lithuania) and Turkey their proportions are higher (above 20%), suggesting ongoing new recruitment to injecting in recent years. Injectors under the age of 25 years account for less than 20% of injectors sampled in 11 countries (10 EU and Turkey), but for over 40% of injectors sampled in Austria, the Czech Republic, Estonia, Latvia, Lithuania, Romania and Slovakia. Most of the countries reporting higher proportions of young injectors experienced the introduction of heroin use later than elsewhere in Europe.

Responses to injecting drug use in Europe have focused on reducing harms such as HIV/AIDS and overdose. Coverage of these measures in the EU has strongly increased since 1995, although it still varies much between countries. The increases in intervention coverage and declines in injecting drug use in the EU appear to be reflected in declines in newly reported HIV cases [3,4].

European countries target injecting drug use and its consequences through a variety of evidence-based interventions [2], mainly in the fields of drug treatment and harm reduction. The most prominent of these is opioid substitution treatment (OST), which is now available in all 27 EU Member States, Croatia and Norway. There are around 650,000 clients in substitution treatment in the EU, representing more than a three-fold increase since 1995. The level of coverage of opioid users, however, shows large variation in the proportion of opioid users with access to OST in ten countries providing estimates: from 5% in Cyprus to over 50% in Germany. Needle and syringe programmes now exist in all 27 EU Member States, Croatia and Norway. Specialised syringe provision outlets — not including pharmacy sales — are estimated to distribute on average about 50 syringes a year per injecting drug user across the EU. In prisons, drug injecting is associated with high levels of syringe sharing, but only five EU countries have implemented needle and syringe programmes in this setting.

It is likely that the declines in injecting drug use and the strongly increasing coverage of key interventions
(including HAART) have resulted in substantial impact on IDU-related HIV transmission in the EU. Rates of newly diagnosed cases of HIV infection among injecting drug users are now mostly at stable and low levels or in decline [4], consistent with the general trend in injecting drug use. The relationship between trends in injecting drug use and IDU-related HIV transmission may be more directly suggested by the example of Bulgaria. Here the drug treatment data indicate an increase in injecting among heroin users in recent years, whereas reported HIV cases among IDUs have also sharply increased [4].

References