South Korea is experiencing the largest outbreak of Middle East respiratory syndrome coronavirus infections outside the Arabian Peninsula, with 166 laboratory-confirmed cases, including 24 deaths up to 19 June 2015. We estimated that the mean incubation period was 6.7 days and the mean serial interval 12.6 days. We found it unlikely that infectiousness precedes symptom onset. Based on currently available data, we predict an overall case fatality risk of 21% (95% credible interval: 14–31).

Primary case

The ongoing outbreak in South Korea began when the primary case developed respiratory illness on 11 May after returning on 4 May from Bahrain (18 April–2 May) via Qatar (2–3 May). Further epidemiological investigation showed that the primary case had also travelled to the United Arab Emirates (29–30 April) and Saudi Arabia (1–2 May) during their stay in Bahrain [2]. Feeling unwell after returning to South Korea, the primary case visited a local clinic (Hospital A) in Pyeongtaek, Gyeonggi province on 12, 14 and 15 May and was hospitalized in Hospital B from 15 to 17 May*. However, this patient did not initially report their recent travel in the Middle East. Upon discharge from Hospital B, the patient visited another clinic (Hospital C) and was admitted to a general hospital (Hospital D) in Seoul on 17 May, where the patient was later diagnosed with MERS-CoV on 20 May. Since then, the patient has been isolated and treated in another hospital designated by the Korean government to treat MERS patients.

Sources of data

We retrieved publicly available data from multiple sources, including the Korea Centers for Disease Control and Prevention (Korea CDC), the Korean Ministry of Health and Welfare (MoH), the WHO and local Korean news reports to compile a line list of all confirmed cases reported by 19 June 2015. In case of any data discrepancy between the different sources, we used the most up-to-date information from official reports published by the Korea CDC and MoH on a daily basis during the outbreak. The official reports were only available in Korean language and included a brief description of each confirmed case, including demographic characteristics (e.g. age and sex), date of exposure and onset of symptoms, as well as possible linkage with confirmed cases and the associated hospital cluster (e.g. Hospital A to P).

Statistical analysis

We fitted parametric distributions to the time intervals (i) from infection to onset (i.e. the incubation period) and (ii) from illness onset to case confirmation. We also fitted a nonparametric distribution on the incubation period. The exact dates of infection were not known for most cases, but exposure windows were available, and we accounted for the consequent interval censoring in the likelihood function [9] and the possibility of infectiousness before illness onset (details on the methodology are available from the corresponding author on request). We used survival models to fit alternative parametric distributions including log-normal, Weibull and gamma distributions, and compared the goodness of fit of these parametric distributions using the Bayesian information criterion. We observed that the delay from illness onset to confirmation shortened as the epidemic progressed, so we fitted two separate survival curves for onset before and after 28 May. We used the same approach to estimate the serial interval.
distribution, based on data on illness onset times for linked cases. We calculated the 95% credible interval (CrI) by bootstrapping.

To estimate the case fatality risk (CFR) allowing for the uncertain clinical outcomes of those who remained in hospital on the date of analysis (19 June 2015), we used the methods proposed by Garske et al. which adjusts the fatality risk based on the time-to-death distribution [10]. We assumed that the time from onset to death followed a log-normal distribution, and used Markov chain Monte Carlo methods to estimate the parameters in a Bayesian framework, setting an informative prior for the time from onset to death with a mean of 14 days [11], and non-informative priors for the other parameters. All statistical analyses were conducted in R version 3.0.2 (R Foundation for Statistical Computing, Vienna, Austria).

**Outbreak description**

The number of laboratory-confirmed cases increased rapidly until 7 June, when 23 cases were confirmed on a single day but appears to have subsided since then (Figure 1A). Figure 1B shows the epidemic curve by date of illness onset for 110 cases with available data. It should be recognised that while the outbreak has not
**Table 1**
Demographic characteristics of confirmed cases of MERS-CoV infection, South Korea, 11 May–19 June 2015 (n = 166)

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>All cases (n=166)</th>
<th>Fatal cases (n=24)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age group</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0–18 years</td>
<td>1 (1%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>19–39 years</td>
<td>31 (19%)</td>
<td>5 (21%)</td>
</tr>
<tr>
<td>40–59 years</td>
<td>64 (39%)</td>
<td>5 (21%)</td>
</tr>
<tr>
<td>60–79 years</td>
<td>61 (37%)</td>
<td>10 (67%)</td>
</tr>
<tr>
<td>≥ 80 years</td>
<td>9 (5%)</td>
<td>3 (13%)</td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>101 (61%)</td>
<td>17 (71%)</td>
</tr>
<tr>
<td>Female</td>
<td>65 (39%)</td>
<td>7 (29%)</td>
</tr>
<tr>
<td><strong>Occupation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Healthcare personnel</td>
<td>30 (18%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Not healthcare personnel</td>
<td>136 (82%)</td>
<td>24 (100%)</td>
</tr>
</tbody>
</table>

MERS-CoV: Middle East respiratory syndrome coronavirus.

yet ended, our preliminary assessment shows that the epidemic to date may have peaked on 1 June when 15 cases reported illness onset. Median age of the 166 cases was 56 years, 101 of 166 (61%) were male, and 30 of 166 (18%) were healthcare personnel (Table 1).

**Transmission chains**

Figure 2 shows a summary sketch of the transmission chain (additional material** showing the detailed chains is available at: http://sph.hku.hk/bcowling/eurosurveillance2015appendix.zip). 119 cases were identified by Korea CDC as having had contact with a confirmed case in the period before their illness onset, and three of these cases had contact with more than one confirmed case. A total of 27 secondary cases in a single hospital have been traced back to the primary case in Hospital B (Figure 2). In particular, Case 14 infected at least 70 cases between 27 and 29 May while being treated in the emergency room in Hospital D, one of the five largest hospitals located in Seoul with 3,980 healthcare professionals and more than 8,000 outpatient visits per day [12]. According to the press conference given at Hospital D on 7 June, at least 893 patients and visitors were potentially exposed to the virus during this period [13], which explains a significant increase in the number of cases confirmed and notified between 6 and 11 June. Since 12 June, when the first fourth-generation case was confirmed, 10 more potential fourth-generation cases have been reported. Because of the marked heterogeneity in transmissibility, with the vast majority of cases associated with just these three superspreading events in the nosocomial setting, it would be misleading to summarily characterise the transmissibility of the virus in this ongoing outbreak with a single average value of the reproductive number [14]. The mean serial interval was 12 to 13 days in each of four epidemiological clusters associated with Cases 1, 14, 15 and 16.

**Epidemiological parameters**

We found that a gamma distribution had the best fit to the incubation period distribution and was very similar to the nonparametric estimate (Figure 3A). The fitted gamma distribution had a median of 6.3 days (95% CI: 5.7–6.8), a mean of 6.7 days (95% CI: 6.1–7.3) and a 95th percentile of 12.1 days (95% CI: 10.9–13.3). Using data on 99 cases with single identified infectors, we found that a gamma distribution with a mean of 12.6 days (95% CI: 12.1–13.1) and standard deviation of 2.8 days (95% CI: 2.4–3.1) provided best fit to the serial interval distribution (Figure 3B). The mean duration of illness onset to laboratory confirmation was 8.1 days for cases with illness onset before May 28, and substantially shorter (mean: 4.4 days) for cases with illness onset after that date (Figure 3C). We used a log-normal regression model for the time from illness onset to laboratory confirmation to estimate that healthcare worker status was not significantly associated with time to confirmation (beta = 0.05; 95% CI: −0.34 to 0.25), with the point estimate signifying a 5% reduction in time to confirmation in healthcare workers.

**Presymptomatic infectiousness**

It appeared that a small number of cases might have been infected before their infectors became symptomatic. Furthermore, Cases 37 and 39 were epidemiologically linked to multiple confirmed cases. To account for the possibility of presymptomatic infectiousness and the uncertainty of who infected Cases 37 and 39 when estimating the incubation period, we (i) simultaneously inferred the incubation period of the infectors of Case 37, (ii) assumed that Case 39 was equally likely to be infected by all cases to whom he had been epidemiologically linked, namely Cases 9, 11, 12 and 14 (because the infecter of Case 39 was not statistically identifiable), and (iii) introduced a parameter $Y$ to represent the time interval between onset of symptoms and onset of infectiousness For example, if cases become infectious two days before onset of symptoms, then $Y$ = 2 days. For a given value of $Y$, the dates of exposure of a case must not precede the date of symptom onset of the case’s infecter by more than $Y$ days. The data were adjusted accordingly during the estimation of the incubation period. Furthermore, we excluded Case 40 when performing the estimation because their exposure and onset date were the same, which was implausible. We used Markov chain Monte Carlo methods to estimate the parameters of this model in a Bayesian framework.
In this modelling analysis of presymptomatic infectiousness, our model suggested that infectiousness might begin 0.4 days (95% CrI: −1.2 to 2.4) before illness onset, which corresponded to a very small (right) shift from the prior distribution. Hence, there was no evidence that infectiousness preceded symptom onset. The same conclusion remained when the standard deviation of the prior was halved or doubled.

Severity of infections
Up to 19 June 2015, 24 cases have died while 30 have recovered and been discharged; the other 112 cases remain in hospital and 16 are in critical condition. Among the 24 fatal cases to date, none of which were in healthcare workers, the median age was 68.5 years (range: 49–83 years). We predicted the final CFR to be 21% (95% CrI: 14–31), allowing for the uncertain outcomes of cases that remained in hospital on the date of analysis.

Comparative epidemiology of MERS and SARS
Table 2 compares key features of the MERS outbreak in South Korea with the features of MERS epidemiology in previous outbreaks in other countries as well as the 2003 outbreak of severe acute respiratory syndrome (SARS) [7,9,11,15-18]. In all MERS outbreaks, current and previous, men were more likely to be cases than women, and the mean age of the cases was around 56 years. There was a marked similarity in the incubation periods and serial intervals and in the case fatality risk.

Discussion
MERS is a relatively new disease, with the first confirmed case reported in Saudi Arabia in 2012 [2,3]. Globally, a total of 1,321 laboratory-confirmed cases of MERS-CoV infection, including 466 deaths, have been reported to the World Health Organization (WHO) to date, of which more than 1,000 occurred in Saudi Arabia [2,4]. One of the major challenges in counting...
the spread of MERS-CoV is the limited understanding of the transmissibility and transmission patterns of the virus, in part because MERS-CoV is a novel pathogen and the experience to date remains mostly confined to cases in Saudi Arabia [4]. However, the outbreak of MERS-CoV in Jeddah, Saudi Arabia in 2014 highlighted an increased transmissibility for secondary human-to-human transmission in healthcare settings [5].

Our findings confirm that the epidemiology of MERS in South Korea is similar to that observed in the Middle East [7] and in fact closely resembles that of the 2002–03 outbreak of SARS [17]. The epidemic thus far has undergone four generations of infection events (Figure 2) arising from delayed recognition of the primary patient who sought care at multiple healthcare facilities before finally being diagnosed and isolated. The Korean outbreak is remarkable in that 148 of 166 transmission events (89%), or 125 of 166 (75%) if those who were epidemiologically linked to a cluster but not any infector are excluded, can be attributed to just three clusters of nosocomial superspreading events (Figure 2). Importantly, there has not been any evidence of community transmission thus far.

Given that (i) there is no known zoonotic reservoir of MERS-CoV in South Korea, (ii) the probability of further foreign importation of infected cases appears to be low because very few MERS cases have been identified outside of the Middle East to date and (iii) infectiousness is unlikely to precede symptom onset, the key to controlling the present epidemic remains prompt recognition and isolation of further cases through rigorous contact tracing and close medical surveillance of those quarantined. This also applies to other outbreaks of MERS that may occur in the future. We estimated that the incubation period had a 95th percentile of 12.1 days, which supports the quarantine period of two weeks currently recommended by public health authorities.

Previous studies based on several outbreaks in the Arabian Peninsula estimated the basic reproductive number ($R_0$) to be between 0.6 and 0.8 overall [6,7,19,20], although with apparent heterogeneity leading to sporadic outbreaks in which $R_0$ exceeded 1 [21]. In our analysis described here we felt that it would not be appropriate to estimate an average reproductive number because of the heterogeneity in transmissibility associated with the three superspreading events. However, it is clear that apart from those three events, the MERS-CoV had low transmissibility in this outbreak.

The CFR of 21% (95% CrI: 14–31) estimated here is substantially lower than the overall CFR in a previous analysis of cases most of whom were from the Middle East (444/1,163; 38%) [2], but the same as the CFR reported by Cauchemez et al. for secondary cases excluding sporadic cases identified after presenting with serious disease (21%) [7], and very similar to the CFR of SARS in Hong Kong in 2003 (Table 2) [17]. While
our estimate of the CFR accounts for uncertainty in the final outcomes of patients that remain in hospital, it is challenging to have accurate estimates of the CFR early in the course of an outbreak [10, 22]. If the CFR in this outbreak remained below 25% once the final outcome for all cases has been ascertained, it would indicate a lower severity of MERS-CoV than in some previous and contemporary reports. A lower CFR would be consistent with the much lower severity observed among secondary cases in the Middle East that were identified through contact tracing, presumably owing to a combination of earlier supportive treatment and/or a lower infective dose and/or lower clinical severity due to other host factors [23]. Greater severity in the sporadic cases could be an artefact of surveillance biased towards infections associated with serious illnesses. Consistent with previous reports, older age was associated with greater risk of severe disease [15, 24, 25]. We did not have data on underlying medical conditions, but it is known from other outbreaks of MERS that a history of chronic disease is another risk factor for disease progression and mortality [11, 15, 25, 26].

Our epidemiological characterisation relied on the assumption that the transmission network as ascertained by the MoH was accurate. Specifically, the network essentially comprised secondary cases of three superspreading events (namely infections caused by Case 1, 14 and 16). The serial interval and incubation period of the secondary cases generated by these three superspreading events were similar, which supports the validity of the network ascertained by the MoH. Nonetheless, infected people with apparently longer incubation periods in the data might have been tertiary instead of secondary cases, in which case we would have overestimated the incubation period. On the other hand, because the outbreak in South Korea is still ongoing and driven by superspreading events, cases with very long incubation periods and/or long serial intervals may not have been identified yet and we may have underestimated the incubation period and serial interval distributions.

This outbreak demonstrates the potential for clusters of emerging infectious diseases to have very substantial societal and economic impact. In South Korea with a population of 50 million, 166 cases of MERS caused major reductions in tourism, nationwide school closures, and some preliminary forecasts for a growth in annual gross domestic product reduced by at least 0.1% [27]. As this outbreak appears to be coming to an end, focus of public health authorities may shift from the immediate control efforts towards a detailed investigation of the mechanisms and causes that led to the superspreading events. The parallels with superspreading events driving the spread of SARS in 2003 in Hong Kong and Singapore emphasise the importance of understanding these events and of determining the measures that could be taken to reduce the risk of similar incidents happening in the future.

** Author’s correction

On request of the authors, the travel dates of the primary case in this sentence were corrected April to May. This change was made on 26 June 2015.

** Note

Additional material made available by the authors on an independent website is not edited by *Eurosurveillance*, and *Eurosurveillance* is not responsible for the content. The material can be accessed at: http://sph.hku.hk/bcowling/eurosurveillance2015appendix.zip.

Acknowledgments

This research was supported by the Harvard Center for Communicable Disease Dynamics from the National Institute of General Medical Sciences (grant no. U54 GM088558), the Health and Medical Research Fund, Food and Health Bureau, Government of the Hong Kong Special Administrative Region (grant no. 14131432), and a commissioned grant from the Health and Medical Research Fund, Food and Health Bureau, Government of the Hong Kong Special Administrative Region (grant no. HKS-15-E05). The funding bodies had no role in study design, data collection and analysis, preparation of the manuscript, or the decision to publish.

Conflict of interest

BJC reports receipt of research funding from MedImmune Inc. and Sanofi Pasteur and consults for Crucell NV. The authors report no other potential conflicts of interest.