Assessment of the Middle East respiratory syndrome coronavirus (MERS-CoV) epidemic in the Middle East and risk of international spread using a novel maximum likelihood analysis approach

C Poletto[1,2,3], C Pelat[3,4,5], D Lévy-Bruhl[6], Y Yazdanpanah[4,5,7], P Y Boëlle[1,2], V Colizza (vittoria.colizza@inserm.fr)[1,2,8]

1. Inserm, U707, Paris, France
2. UPMC Université Paris 06, Faculté de Médecine Pierre et Marie Curie, UMR S 707, Paris, France
3. These authors contributed equally to this work
4. ATIP/Avenir Inserm, U738, Paris, France
5. Univ Paris 07, Faculté de Médecine Bichat, Paris, France
6. Institut de Veille Sanitaire (InVS), St Maurice Cedex, France
7. Service des Maladies Infectieuses et Tropicales, Hôpital Bichat Claude Bernard, Paris, France
8. Institute for Scientific Interchange (ISI), Turin, Italy

The emergence of the novel Middle East (ME) respiratory syndrome coronavirus (MERS-CoV) has raised global public health concerns regarding the current situation and its future evolution. Here we propose an integrative maximum likelihood analysis of both cluster data in the ME and importations in a set of European countries to assess the transmission scenario and incidence of sporadic infections. Our approach is based on a spatial-transmission model integrating mobility data worldwide and allows for variations in the zoonotic/environmental transmission and under-ascertainment. Maximum likelihood estimates for the ME, considering outbreak data up to 31 August 2013, indicate the occurrence of a subcritical epidemic with a reproductive number R of 0.50 (95% confidence interval (CI): 0.30–0.77) associated with a daily rate of sporadic introductions p_{sp} of 0.28 (95% CI: 0.12–0.85). Infections in the ME appear to be mainly dominated by zoonotic/environmental transmissions, with possible under-ascertainment (ratio of estimated to observed (0.116) sporadic cases equal to 2.41, 95% CI: 1.03–7.32). No time evolution of the situation emerges. Analyses of flight passenger data from ME countries indicate areas at high risk of importation. While dismissing an immediate threat for global health security, this analysis provides a baseline scenario for future reference and updates, suggests reinforced surveillance to limit under-ascertainment, and calls for alertness in high importation risk areas worldwide.

Introduction

As of 31 August 2013, a total of 108 laboratory-confirmed cases of human infection with the Middle East respiratory syndrome coronavirus (MERS-CoV) have been reported to the World Health Organization (WHO) [1]. Since the first identification of the virus in 2012 [2], a rapid coordinated response has been put in place to confront the novel emerging epidemic. This response consisted in the enhancement of surveillance systems, the provision of updated information on the epidemic situation, technical guidance for the clinical management of probable infections [3-7], and the search for the possible virus reservoir [8,9]. There are still many uncertainties about various aspects of the outbreak, including its full geographical extent, a possible extension of an initial virus reservoir to other hosts, the transmission path of the infection to humans and the associated risk. All these aspects call for heightened surveillance, enhanced investigations and the development and application of epidemiological methods to assess the epidemic situation and determine the potential of the virus to spread in humans and to circulate at a global scale.

In such a situation, statistical, mathematical and computational methods allow estimating key epidemiological parameters from available data, under various assumptions and accounting for the many uncertainties. The reproductive number R, i.e. the average number of secondary cases generated by a primary case, is a key summary measure of the transmissibility of an emerging infection. A first estimation of the MERS-CoV reproductive number was based on the analysis of cluster-size data with assumed cluster partition in terms of transmission trees, highlighting the similarity of the current MERS-CoV situation to the pre-epidemic stage of the severe acute respiratory syndrome (SARS) outbreak [10]. Aside from the

Number of clusters of a given size, depending on the total number of cases of Middle East respiratory syndrome coronavirus considered

Size of cluster (number of cases)	Number of clusters[a]			
	Baseline (all laboratory-confirmed cases in the ME region[b] as of 31 August 2013)	All laboratory-confirmed cases in the ME region[b] as of 31 August 2013, and the probable cases in the Jordan April 2012 outbreak	All laboratory-confirmed cases worldwide as of 31 August 2013	All laboratory confirmed-cases in ME region[b] as of 31 May 2013
1[c]	42	42	44	16
2	8	7	10	2
3	2	2	4	1
5	2	2	2	0
10	0	1	0	0
22	1	1	1	1
Total number of cases	96	104	108	45

ME: Middle East.

[a] Unless otherwise specified.
[b] Including Jordan, Qatar, Saudi Arabia, and United Arab Emirates.
[c] Sporadic case.

transmissibility, an additional important characteristic of the epidemic remains unknown i.e. the incidence of infection. Observed cases may indeed only represent a proportion of the current epidemic, with a majority of infections going undetected because of mild illness or asymptomatic infection. This aspect also has further relevant implications for the correct estimation of other important overall statistics (e.g. the severity of the disease) and of the risk of importation of cases from affected areas to other locations of the world. Limited data may also hide important changes in the virus transmissibility related, for example, to viral adaptations to humans that may alter its pandemic potential, thus presenting an additional challenge for the assessment of the epidemic situation.

To fill the gaps in current knowledge, we present here an innovative integrative maximum likelihood approach to describe the epidemic in the Middle East (ME) region, comprising Jordan, Qatar, Saudi Arabia and the United Arab Emirates. We synthesise evidence from multiple sources of information: sizes of clusters of cases, traffic data, and imported cases outside the region. The methods used account for the limited information available and reporting inaccuracies. Our aim is to complete early findings on the MERS-CoV epidemic by focusing on the virus transmissibility from human-to-human, its possible changes in time, and the expected number of cases in the ME region. Based on international travel flows, the public health threat for other geographical locations not yet affected by the virus is also assessed.

Methods

Analytic overview

The integrative approach we use is based on a combined maximum likelihood analysis to jointly estimate the reproductive number R and the daily rate p_{sp} of sporadic introduction of the virus in the population through zoonotic/environmental transmissions [11]. The integrative approach builds on two aspects of the currently reported outbreak – the distribution of cluster sizes, providing information on R (Method 1), and the number of imported cases in countries out of the source region providing information on R and p_{sp} based on the fit of a stochastic spatial metapopulation model integrating aviation data worldwide (Method 2).

Method 1

We considered laboratory-confirmed cases reported to WHO as of 31 August 2013 in the ME region including Jordan, Qatar, Saudi Arabia, and United Arab Emirates [1] (Table 1). This region is also referred to as the source region in this study. The cases were distributed into clusters according to information found in WHO reports. Using this dataset as baseline data, we estimated the reproductive number R using the cluster sizes distribution. Several distributions can be used that correspond to different hypotheses regarding the number of secondary cases distribution (offspring distribution) [12]. In particular we considered a Poisson offspring distribution accounting for no overdispersion around a common mean [13] and a geometric offspring distribution assuming a constant rate of transmission during an exponentially distributed infectious period [14].

As cluster sizes may be biased downwards by incomplete observation, we allowed for uncertainty by assuming that each case in a cluster would go unobserved with probability p_{cl} during investigation ($p_{cl}=0$ representing no missed cases). This corresponds to the following distribution for reported cluster sizes:

$$P(O = k|R, p_{cl}, O \geq 1) = \frac{\sum_{j \geq k} P(S=j|R)\binom{j}{k}p_{cl}^{j-k}(1-p_{cl})^k}{1-P(O=0|R,p_{cl})}, \quad (1)$$

where O is the observed size of the cluster, S its real size

$(O \leq S)$, $P(O = 0|R, p_{cl}) = \sum_{j \geq 1} P(S = j|R)p_{cl}^{j}$, and $P(S|R)$

is the offspring distribution discussed above. Eventually, the likelihood was computed as

$$\mathcal{L}_1(R, p_{cl}; \{o_i\}) = \left(\sum_i o_i\right)! \prod_i \frac{1}{o_i!} P(O_i = o_i|R, p_{cl}, O \geq 1)^{o_i}$$

over a bi-dimensional grid of $\{R, p_{cl}\}$ values.

We performed a sensitivity analysis by considering: (i) the addition, to the baseline data, of the complete Jordan cluster including eight more cases (cluster size=10, Table 1) identified through a retrospective serology study carried out on 124 individuals [4,15]; (ii) all laboratory-confirmed cases (n=108) reported worldwide to WHO as of 31 August 2013 [1]; (iii) laboratory-confirmed cases in the ME region up to 31 May, 2013 [1]. The corresponding cluster size data are reported in Table 1.

Method 2
Due to a large concern around the ongoing outbreak and enhanced surveillance following the WHO guidelines for patients returning from the affected area, the detection of probable cases imported in countries out of the ME region is expected to be more complete than in the region itself where primary cases may have gone undetected. Another source of information to estimate the reproductive number R, discounting possible notification/surveillance biases in the source region, is therefore provided by the importation of cases in newly affected countries [16,17]. As a basis for our estimation model, we use a method already employed for the estimation of the seasonal transmission potential of the 2009 influenza A(H1N1)pdm09 virus, which was based on the calibration of a global epidemic and mobility model (GLEAM) [18,19] to the chronology data of the 2009 pandemic [17]. We modified and extended this method by accounting for the different transmission scenarios in the ME region. The method concurrently allows the estimation of the incidence of infection from sporadic cases in the region, and therefore provides a measure of possible under-ascertainment of cases.

GLEAM is based on a spatially structured metapopulation approach comprising 3,362 subpopulations in 220 countries in the world coupled through mobility connections. The model is informed with high-resolution demographic data for six billion individuals and multiscale mobility data including the full air traffic database from the International Air Transport Association (IATA) and short-range ground mobility obtained from national commuting data [19]. The infection dynamics takes place within each subpopulation and assumes a modified susceptible, exposed, infectious, recovered individuals (SEIR) compartmentalisation [20] to account for different transmission scenarios in the ME region [11]: (i) introduction of sporadic infections from zoonotic/environmental transmission with a daily rate p_{sp} and a uniform spatial distribution in the source region; (ii) modified human-to-human transmissibility with respect to standard homogeneous mixing in all subpopulations of the model to allow for large variations in the number of secondary cases produced by a given primary case. Epidemiological parameters for the compartmental model were based on the estimates obtained from the analysis of the outbreak data including 22 cases at a healthcare facility in Al-Ahsa in Saudi Arabia [6], namely average latency period of 5.2 days and generation time of 7.6 days.

The daily rate p_{sp} of sporadic cases emergence in the ME region and the reproductive number R are the free parameters of the model. For each set of values of these two parameters, GLEAM allows the generation of stochastic numerical realisations of the MERS-CoV outbreak simulating the local epidemic in the source region and the possibility of international dissemination through mobility processes entirely based on real data. We thus generate with a Monte Carlo procedure the probability distribution $P_i(n_i)$ of the number n_i of imported MERS-CoV cases in country i out of the source region as of 31 August 2013 (4×10^3 stochastic realisations for each point (R, p_{sp}) of the space of parameters). Being all independent importation events, we can define a likelihood function

$$\mathcal{L}_2(R, p_{sp}; \{n_j^*\}) = \prod_j P_j(n_j^*)$$

where n_j^* is the empirically observed number of imported cases per country (see schematic example in Figure 1). We further restrict our analysis to certain western European countries including Austria, Belgium, Denmark, Finland, France, Germany, Iceland, Ireland, Italy, Liechtenstein, Luxemburg, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, and the United Kingdom, to focus on an area where respiratory diseases surveillance is homogeneous and with a high sensitivity to detect importations. We disregard other possible sources of heterogeneity. These data consisted therefore in: $n_j^*=1$ for France, Germany, Italy and the United Kingdom, and $n_j^*=0$ for all other western European countries j considered in the analysis (Figure 1). We estimated the log-likelihood over a bi-dimensional grid of (R, p_{sp}) values and used bivariate linear interpolation over a refined grid.

FIGURE 1

Maximum likelihood approach integrating two methods, to estimate the daily rate of sporadic cases of Middle East respiratory syndrome coronavirus infection and the reproductive number
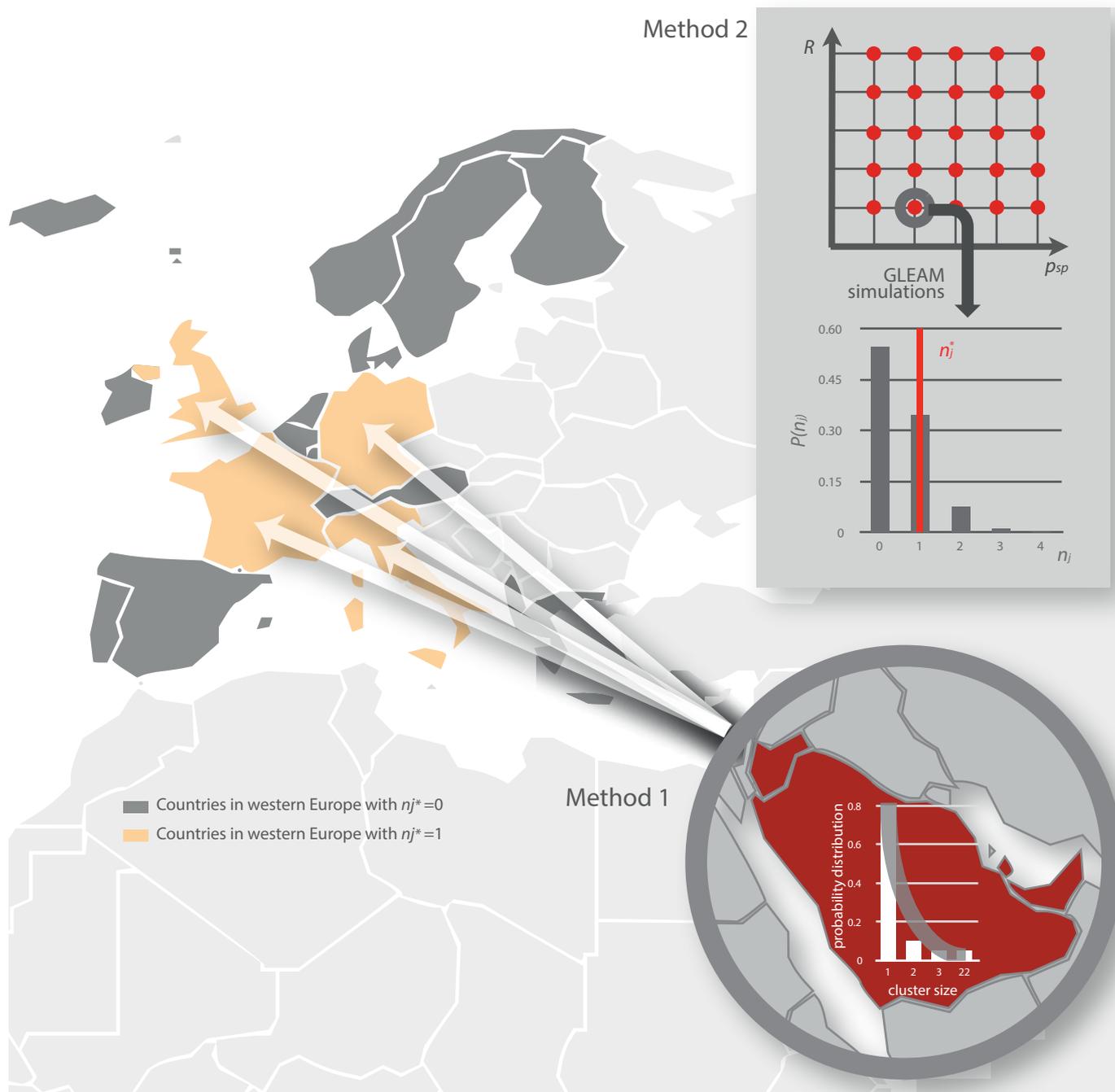
GLEAM: global epidemic and mobility model; p_{sp}: daily rate of sporadic introduction of the virus in the population through zoonotic/environmental transmissions; R: reproductive number.

In this maximum likelihood approach, the source region used comprises Jordan, Qatar, Saudi Arabia, and United Arab Emirates.

Method 1 (bottom circle) is based on the maximum likelihood analysis of cluster size distribution obtained from laboratory-confirmed cases in the source region (countries in red in the zoomed area).

Method 2 (top panel) is based on the maximum likelihood analysis on data on case importations in certain western European countries, as schematically indicated on the map (countries included are: Austria, Belgium, Denmark , Finland, France, Germany, Iceland, Ireland, Italy, Liechtenstein, Luxemburg, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, and the United Kingdom). For each point in the parameter space (R, p_{sp}) we run 4,000 stochastic GLEAM simulations from the same initial conditions and parameterised as described in the main text. With each run providing the simulated number of imported cases n_j for a given country j, we can compare the resulting simulated probability distribution of n_j with the observed value n_j* for that country as of 31 August 2013, and compute a likelihood function for all western European countries included in the analysis.

Table 2

Best estimate values for the reproductive number R and the daily rate p_{sp} of emergence of sporadic cases of Middle East respiratory syndrome coronavirus due to zoonotic or environmental transmissions

Analysis	Data	R (95% CI)	p_{sp} (95% CI)
Baseline	Baseline data: all laboratory-confirmed cases in the ME region[a] up to 31 Aug 2013	0.50 (0.30–0.77)	0.28 (0.12–0.85)
Sensitivity	All laboratory-confirmed cases in the ME region[b] as of 31 Aug 2013, and the probable cases in the Jordan Apr 2012 outbreak	0.65 (0.34–0.80)	0.28 (0.12–0.83)
	All laboratory-confirmed cases reported worldwide to WHO as of 31 Aug 2013	0.50 (0.31–0.77)	0.28 (0.12–0.85)
	All laboratory-confirmed cases reported in an extended source region including ME region and neighbouring countries[b]	0.50 (0.31–0.76)	0.14 (0.05–0.38)[c]
	Data considering a restricted source region limited to Saudi Arabia	0.60 (0.30–0.76)	4.73 (2.32–15.37)[c]
	Data considering all laboratory-confirmed cases in ME region[a] as of 31 May 2013	0.54 (0.34–0.90)	0.43 (0.12–0.95)
Alternative baseline scenario	Baseline data considering a Poisson offspring distribution	0.69 (0.34–0.79)	0.28 (0.12–0.71)

CI: confidence interval; ME: Middle East; WHO: World Health Organization.

Results for the baseline and for the scenarios of the sensitivity analysis are obtained assuming a geometric offspring distribution for the analysis of cluster data and for the best estimate of the uncertainty parameter, p_{cl}. The last row of the Table refers to the baseline scenario considering a Poisson offspring distribution for the analysis of cluster data.

[a] Jordan, Qatar, Saudi Arabia and the United Arab Emirates.
[b] Bahrain, Iraq, Iran, Israel, Kuwait, Lebanon, Oman, Palestine, Syria, and Yemen in addition to the ME region.
[c] To be comparable with the other estimates, this value has been rescaled to take into account the change of population size of the source region; it thus represents the daily rate of sporadic cases scaled to the ME region.

Given the unknown geographical extension of the source of the MERS-CoV infection in the ME and its reservoir, we performed a sensitivity analysis by considering: (i) an extended definition of source region to neighbouring countries, thus additionally including Bahrain, Iraq, Iran, Israel, Kuwait, Lebanon, Oman, Palestine, Syria, and Yemen (extended source region), based on travel recommendations [21]; (ii) a restricted source region localised only in Saudi Arabia, i.e. the country in the ME that reported the largest number of cases; (iii) a variation in the time of the initial emergence of the virus with transmission to humans, assuming that sporadic cases may be introduced up to two months before the known initial cases (Jordan cluster, April 2012 [4]), to allow for lack of identification or detection prior to the Jordan cluster.

Integrative approach combining Methods 1 and 2

Methods 1 and 2 were jointly combined in the following integrated likelihood function:

$$\mathcal{L}(R, p_{sp}, p_{cl}|\{o_i\},\{n_j^*\}) = \mathcal{L}_1(R, p_{cl}|\{o_i\})\mathcal{L}_2(R, p_{sp}|\{n_j^*\})$$

owing to the independence of the two observed processes (cluster sizes and importations). Maximum likelihood estimates were computed over the three-dimensional (3D) grid (R, p_{sp}, p_{cl}) and the deviance

$$D(R, p_{sp}) = -2\big(\log\mathcal{L}(R, p_{sp}) - \max(\log\mathcal{L})\big)$$

was used to measure distance from the best fit. Associated confidence intervals were obtained by profiling the deviance in the 3D space [22]. It is important to note that such estimates cannot be derived from the maximum likelihood analysis of each Method considered separately, nor conditionally one to the other, and the full computation of

$$\mathcal{L}(R, p_{sp}, p_{cl}|\{o_i\},\{n_j^*\})$$

needs to be considered. In this respect, our integrative approach represents a substantial advance with respect to prior work based on the analysis of cluster data only [10].

FIGURE 2

Air traffic capacity in the Middle East respiratory syndrome coronavirus source region considered in this study and international destinations from this region
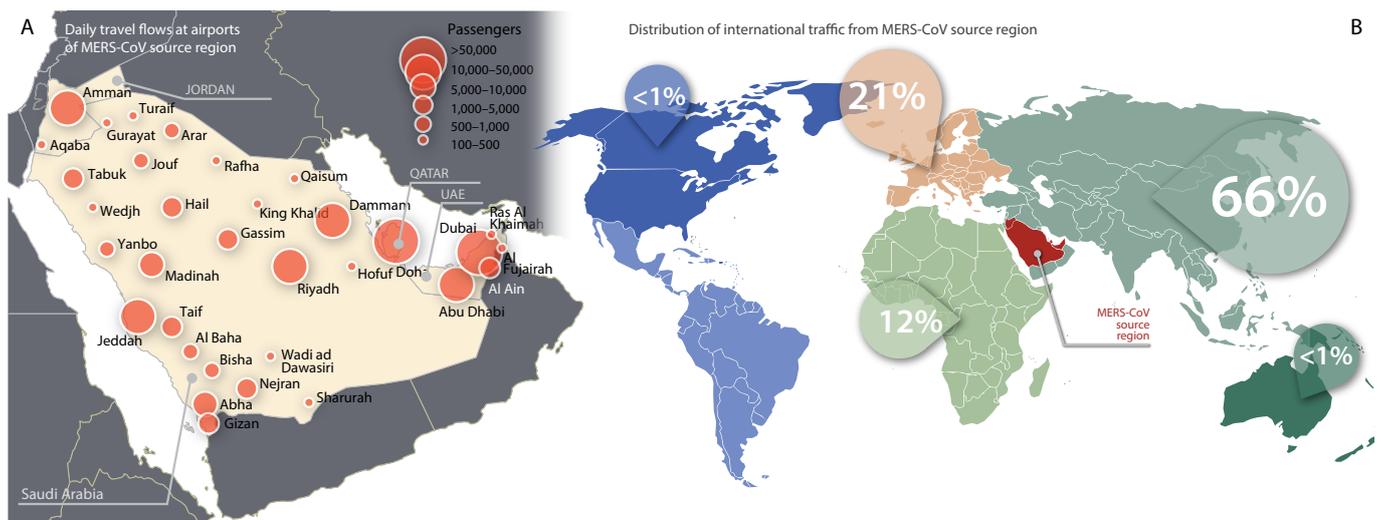
MERS-CoV: Middle East respiratory syndrome coronavirus; UAE: United Arab Emirates.

Airports in the MERS-CoV source region (Jordan, Qatar, Saudi Arabia, UAE) are represented with a circle proportional to the daily traffic they handle (panel A). Their international traffic, out of the region, is broken down by continent of destination (panel B). Breakdown by country for the first 20 countries with highest traffic from the ME region: India (11.7%), Bahrain (8.7%), Pakistan (8.6%), United Kingdom (8.4%), Oman (5.8%), Egypt (5.2%), Kuwait (4.3%), Iran (3.6%), Germany (3.5%), Lebanon (2.9%), Bangladesh (2.8%), Thailand (2.5%), Sri Lanka (2.3%), Singapore (2.1%), France (2.0%), Syria (2.0%), Kenya (1.6%), Italy (1.5%), Malaysia (1.4%), Switzerland (1.4%). Statistics are based on the 2002 International Air Transport Association (IATA) air traffic data for direct flights integrated into the global epidemic and mobility model (GLEAM) [19] after accounting for traffic growth in the period from 2002 to 2011 [23].

Air traffic data analysis

We additionally analysed the air traffic data integrated into GLEAM (direct flights from reference [19] accounting for traffic growth [23]) to evaluate the traffic capacity of the airports in the ME region and to assess the importation risk of the countries belonging to other areas of the world than the western European countries previously considered.

Results

The integrated analysis based on outbreak data up to 31 August 2013 led to a R value equal to 0.50 (95% CI: 0.30–0.77) and daily rate p_{sp} of MERS-CoV introductions into the human population in the ME region equal to 0.28 (95% CI: 0.12–0.85) (Table 2). These best estimates were obtained considering a geometric offspring distribution, yielding higher maximum likelihood values in the analysis. The corresponding best estimate for uncertainty in the cluster distribution suggests a consistent fraction of cases missed in cluster investigations (p_{cl}=0.35; 95% CI: 0–0.85), but little impact of inaccuracies in reported cluster size on the estimates of the other parameters.

The estimated daily rate of sporadic cases in the ME region (p_{sp}=0.28; 95% CI: 0.12–0.85) can be compared to the observed value ($p^{\star}_{sp}$=0.116), computed based on 60 sporadic cases reported between April 2012 and 31 August 2013, including 42 sporadic cases and

13 cluster index cases in the ME region (assuming that each cluster is originated by a single index case), as well as five laboratory-confirmed cases exported from the region (four to western Europe [1,3,5] and one to Tunisia [1]). This yields that the true number of cases from zoonotic/environmental transmission might be between 1.03- and 7.32-fold the observed number.

Our estimates are very robust against the addition of cases out of the ME region to the distribution of clusters sizes (R=0.50; 95% CI: 0.31–0.77; no change for p_{sp}). An increase in R, though with limited change in the associated confidence interval, is obtained if we include the full Jordan cluster of April 2012 by considering also cases retrospectively confirmed by serology (R=0.65; 95% CI: 0.34–0.80); no variations are obtained in the confidence interval of the estimated daily rate of sporadic cases in the region (Table 2).

Similar results for the reproductive number are obtained when we consider variations in the geographical definition of the MERS-CoV source region. Extending the source region considered here to neighbouring countries does not affect the estimated basic reproductive number and associated confidence interval (R=0.50; 95% CI: 0.31–0.76), but lowers the value of the daily rate of sporadic cases (p_{sp}=0.14; 95% CI: 0.05–0.38). If we assume that the source region is instead restricted to Saudi Arabia, a substantial increase in p_{sp} is obtained (4.73; 95% CI: 2.32–15.37), with increase

FIGURE 3

Heatmap of deviance values versus reproductive number R and daily rate of sporadic cases p_{sp}
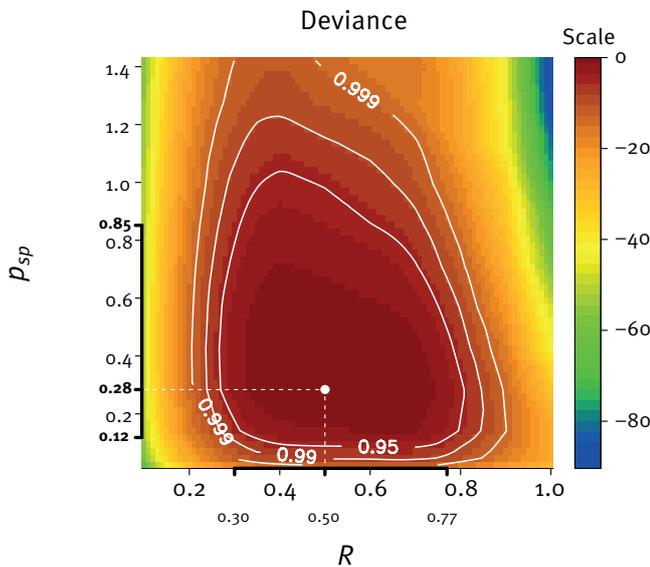
Deviance was calculated as $D(R,p_{sp})=-2(\log\mathcal{L}(R,p_{sp})-\max(\log\mathcal{L}))$ using the profiled log likelihood (for each pair (R,p_{sp}), the uncertainty parameter p_{cl} in cluster size distribution maximizing the log-likelihood $\mathcal{L}$ was chosen). Vertical and horizontal dashed lines show the maximum likelihood values for R and p_{sp}, respectively. Solid white curves contour the deviance-based confidence regions of levels 95%, 99%, and 99.9%. The 95% profiled confidence intervals for R and p_{sp} are highlighted in bold on the axes.

in the reproductive number and unaltered confidence interval ($R=0.60$; 95% CI: 0.30–0.76).

No variation in p_{sp} was observed in testing a different hypothesis on the offspring distribution, whereas an increase for the best estimate of R was found (0.69 when considering a Poisson offspring distribution vs. 0.50 in the baseline, with similar CIs, Table 2). Larger CIs, but no significant variation in the parameters' estimates, were observed by considering empirical data up to the end of May (Table 2).

Analyses of traffic data expose large traffic fluxes towards the continents of Asia, Europe and Africa (Figure 2) from the ME region. Of the 20 countries with highest traffic from the ME region, six were found to neighbour this region (Bahrain, Iran, Kuwait, Lebanon, Oman and Syria), seven were identified in south Asia (Bangladesh, India, Malaysia, Pakistan, Singapore, Sri Lanka, Thailand) and five, including Switzerland, were in Europe, and these comprised the four countries (France, Germany, Italy and the United Kingdom) reporting importation of cases from the affected area. Two countries (Egypt and Kenya) were also found in Africa.

Discussion

Results of our integrative modeling approach suggest the occurrence of a subcritical MERS-CoV epidemic in the ME region, as quantified by a reproductive number smaller than one. The outbreak is not able to generate a self-sustaining epidemic in humans, and sporadic cases from zoonotic/environmental transmission are expected to represent a large fraction of the total size of the epidemic.

The estimated CI for the reproductive number is found to be very stable across changes in the data interpretation. In all cases, considering data up to 31 August 2013, we found that it is highly unlikely (<5% probability) to have a MERS-CoV outbreak with R above 0.80 or below 0.30. The variation of the best estimate from the baseline case ($R=0.50$) to the various scenarios explored as sensitivity analysis (up to $R=0.69$) is explained by the presence of a large region in the parameter space (R, p_{sp}) where the likelihood function shows small variation around its maximum value (darker red area in Figure 3). This is likely induced by the limited data available not allowing us to further narrow down the confidence intervals of the estimates.

The analysis based on the integration of two independent methods allows us to provide an estimate for the daily rate of introductions of MERS-CoV infections in the human population in the ME region, in addition to the estimate for the reproductive number. The estimated 95% CI in the baseline scenario (0.12–0.85) compared to the observed value (0.116) suggests a negligible to significant under-ascertainment rate for zoonotic/environmental transmissions (1.03–7.32 times the reported sporadic cases), indicating that notified sporadic cases likely represent a substantial proportion of the total, but improved surveillance in the region including serological surveys around cases is needed. Since evidence for mild illness, as well as for a wide spectrum of clinical disease, was observed [3,6], our findings are compatible with an under-ascertainment rate for zoonotic/environmental transmissions that may be due in a large part to a selection bias towards cases of more severe illness, where patients having mild illnesses or asymptomatic infections may go undetected [24].

The integrative approach allows us to overcome scarce data availability on the outbreak that may limit the statistical power of each of the approaches if considered separately. The combination of the likelihood functions indeed enables solving possible degeneracies and providing point estimates and confidence intervals for both parameters, R and p_{sp}.

Our estimates for the reproductive number are consistent with the results of Breban et al. [10] – the only study to date reporting results on inter-human transmissibility – thus further confirming the robustness of our epidemic assessment. Our work presents however substantial differences in the methodology and in its achievable predictions as further discussed below.

One major difference is that our integrative approach allows the quantification of sporadic cases under-ascertainment through the estimate of the rate of introduction of sporadic cases in the ME region, combined with the estimate for the reproductive number. In the study from Breban et al. [10], daily introductions are calculated on the basis of the two assumed scenarios for the transmission trees, i.e. from the assumed number of index cases among the reported data. Our procedure instead makes no assumption on the completeness of reported data, or on the local transmission trees, and relies on alternative data sources (case importations) to estimate the number of sporadic cases in the region.

The cluster data analysis of Method 1 relies on the assumption that the final size of the cluster is observed and that each cluster is the result of human-to-human transmission starting from a single index case. The first hypothesis was almost met, since, as of 7 September 2013, only two cases reported after 31 August 2013 were later linked to the cases analysed here, with no impact on the reported estimates [25].

Considering the second hypothesis, while we allow for uncertainty in case detection in the close contact investigation, we do not consider the possibility of co-exposure of epidemiologically linked cases to the same source of zoonotic/environmental infection, differently from Breban et al. [10]. In view of the persisting large uncertainties regarding the virus path of infection to humans and with insufficient data from epidemiological investigations to reliably reconstruct transmission trees within clusters, we chose a worst-case assumption for the transmissibility of the virus. This may lead to overestimating the reproductive number, however not affecting our conclusion on the subcritical nature of the current MERS-CoV epidemic. In addition, such assumption does not affect the estimate of the size of the epidemic in the affected region (Method 2), as this is based on case importations assuming no knowledge on the local transmissions, independently of their type (whether human-to-human or zoonotic/environmental).

We considered the cluster analysis (Method 1) restricted to the ME region as we assumed a rather homogeneous implementation of control measures around cases that may be different from the one put in place in affected countries experiencing importation of cases, mainly due to the additional available knowledge of travel history associated to imported cases. The extension of the analysis to all MERS-CoV clusters of laboratory-confirmed cases reported to WHO did not alter our estimates. The integrative approach also relies on the assumptions of homogeneous mixing in the local populations and homogeneous travel behaviour informed by traffic data (regardless e.g. of travel frequencies of specific population classes), similarly to previous studies [16,17]. This is due to the lack of data characterising interactions between travellers and local population, and characterising the demographic profile of passengers.

The spatial component of our approach allows us to shed light on additional aspects of the epidemic. If we assume that the source of MERS-CoV infection is restricted to Saudi Arabia, where the majority of cases has been observed, our estimates indicate that a much larger number of sporadic cases in the area would be needed to sustain the observed importation of cases in the western European countries considered (4.73 daily introductions of sporadic cases vs. 0.28 in the baseline). The biggest airports, handling the vast majority of the international air traffic of the region, are indeed mainly localised in the United Arab Emirates, Qatar, and Jordan (panel A of Figure 3). This strongly reduces the traffic capacity of the ME region, as well as the corresponding likelihood of exporting cases out of the region, when the restricted hypothesis on the source region is considered. An analysis based on cluster data only would remain unchanged and would not be able to detect important variations in the epidemic size estimation. This result further indicates the relevance of air travel in the epidemic assessment and therefore the need for an integrative approach also based on mobility and space. It also implies that improving the knowledge of the geographical extent of the source region is critical, along with the identification of the virus' path of transmission to humans.

These results are obtained under the assumption of homogeneous sensitivity of the surveillance systems of the restricted set of countries in western Europe with case importations considered in Method 2. The further application of this approach to a wider range of countries or other outbreaks would require the assessment of the sensitivity of the surveillance systems of the affected countries.

Changes in time of countries' public health actions for surveillance and control of the MERS-CoV epidemic certainly occurred in response to the increasingly available emerging evidence and the higher awareness of the disease, however data are too scarce to provide estimates of the reproductive number as a function of time R(t). Here we assumed a constant R for the period under study, with the underlying assumption of a constant and homogeneous implementation of intervention measures in the region.

Other possible factors leading to variations of the current situation may be pathogen- and host-related. Evolution of the MERS-CoV virus to adapt to humans and reach sustainable and efficient human-to-human transmission represents a potential future scenario, as it happened for SARS [26]. Seasonality may also affect virus transmissibility. We tested for possible variations of R and p_{sp} estimates by comparing two different points in time (end of May and end of August 2013) and found no significant change, except for a reduction of the CIs following a larger dataset available for the estimation.

Given the available data, this result seems to indicate that no variation in the rate of introductions and in the transmissibility of the virus has occurred in the period from May to August 2013 that may point to differences in the transmission to humans or to viral adaptations to human hosts. Furthermore, the crude notification data from the affected countries do not provide any epidemiological evidence of a seasonality of the epidemic. On the other hand, if compared to the reported number of sporadic cases at the two dates, our p_{sp} estimates are in favour of an increase in sporadic case ascertainment as the estimated under-ascertainment rate has decreased from 2.2–17.5 considering cases up to 31 May to 1.03–7.32 up to 31 August. Further data to update the integrative approach will contribute to provide a continuous assessment of the outbreak in case an evolving situation is suspected.

Other events that are related to human movements and mixing may as well alter the assessed scenario. Vast international mass gatherings annually taking place in Saudi Arabia are known to bring large number of pilgrims to the affected area, with expected increased rates of local mixing that may favour the transmission of the virus, followed by a potential amplification of its international dissemination due to the return of pilgrims to their own countries [27]. The occurrence of these large scale events calls for additional studies in pilgrims' screening [28] next to enhanced local surveillance in the region and guidance to local authorities [29] that would help to assess and control possible changes in time in the virus transmission.

Air travel clearly represents the main mean for global spatial spread of infectious disease epidemics in the modern world, as it was previously experienced with SARS and the influenza A(H1N1)pdm09 pandemic [16,17,30-32]. Besides seasonal variations due to specific events (e.g. mass gatherings) or in/out flows of expatriates for seasonal jobs, a potential emerging pandemic in the ME area would constitute a very high risk for considerably rapid and wide international spread. The ME area indeed covers a central role in connecting different regions of the world and has witnessed a dramatic increase of traffic growth in the last decade (153% of relative increase in Saudi Arabia in the period from 2002 to 2011, 240% in Jordan, 408% in Qatar, 512% in the United Arab Emirates, against a global relative increase of 168%) [23]. Analysis of the air traffic data integrated into GLEAM suggests that other countries than the ones already affected are at high risk of MERS-CoV importation through infected passengers, in particular in the southern regions of Asia (Figure 2), similarly to reference [27]. Should the outbreak evolve in a self-sustained epidemic, such risk assessment analyses cannot rely on travel data only and would require the full integration of the air travel data with an epidemic model, as in GLEAM, to explicitly simulate the evolving epidemic, estimate importation likelihood [31] and provide predictions for future stages of the epidemic [17,33].

With a subcritical epidemic in the ME region associated with a large potential for international dissemination, priority for the epidemic control should be given to the identification of the transmission of infection to humans to limit sporadic cases, to the reduction of human-to-human transmission through rapid case identification and isolation, and to the enhancement of surveillance systems in those countries that are at a higher risk of importation because of travel flows to/from the affected area.

Conflict of interest

None declared.

Authors' contributions

CaP and PYB did the analyses for Method 1. ChP and VC did the analyses for Method 2. ChP devised an updated version of GLEAM adapted to the MERS-CoV study and ran GLEAM simulations. ChP and CaP performed the integrative analysis. PYB and VC conceived and designed the study. All authors were involved in the interpretation of results and the preparation of the manuscript. VC drafted the manuscript. All authors provided comments and approved the final version of the manuscript.

References

1. World Health Organization (WHO), Global Alert and Response (GAR), Coronavirus infections. Geneva: WHO. [Accessed 11 Sep 2013]. Available from: http://www.who.int/csr/disease/coronavirus_infections/en/index.html

2. Zaki AM, van Boheemen S, Bestebroer TM, Osterhaus AD, Fouchier RA. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. N Engl J Med. 2012;367(19):1814-20. http://dx.doi.org/10.1056/NEJMoa1211721

3. Health Protection Agency (HPA) UK Novel Coronavirus Investigation team. Evidence of person-to-person transmission within a family cluster of novel coronavirus infections, United Kingdom, February 2013. Euro Surveill. 2013;18(11). pii: 20427.

4. Hijawi B, Abdallat M, Sayaydeh A, Alqasrawi S, Haddadin A, Jaarour N, et al. Novel coronavirus infections in Jordan, April 2012: epidemiological findings from a retrospective investigation. East Mediterr Health J. 2013; 19: S12-8.

5. Guery B, Poissy J, el Mansouf L, Sejourne C, Ettahar N, Lemaire X, et al. Clinical features and viral diagnosis of two cases of infection with Middle East respiratory syndrome coronavirus: a report of nosocomial transmission. Lancet. 2013; 381(9885):2265-72. http://dx.doi.org/10.1016/S0140-6736(13)60982-4

6. Assiri A, McGeer A, Perl TM, Price CS, Al Rabeeah AA, Cummings DA et al. Hospital outbreak of Middle East respiratory syndrome coronavirus. N Engl J Med. 2013; 369(5):407-16. http://dx.doi.org/10.1056/NEJMoa1306742

7. Buchholz U, Müller MA, Nitsche A, Sanewski A, Wevering N, Bauer-Balci T, et al. Contact investigation of a case of human novel coronavirus infection treated in a German hospital, October-November 2012. Euro Surveill. 2013;18(8). pii:20406.

8. Memish ZA, Mishra N, Olival KJ, Fagbo SF, Kapoor V, Epstein JH, et al. Middle East respiratory syndrome coronavirus in bats, Saudi Arabia. Emerg Infect Dis. 2013;19(11):1819-23. http://dx.doi.org/10.3201/eid1911.131172

9. Perera RA, Wang P, Gomaa MR, El-Shesheny R, Kandeil A, Bagato O, et al. Seroepidemiology for MERS coronavirus using microneutralisation and pseudoparticle virus neutralisation assays reveal a high prevalence of antibody in dromedary camels in Egypt, June 2013. Euro Surveill. 2013; 18(36). pii:20574.

10. Breban R, Riou J, Fontanet A. Interhuman transmissibility of Middle East respiratory syndrome coronavirus: estimation of pandemic risk. Lancet. 2013;382(9893):694-9. http://dx.doi.org/10.1016/S0140-6736(13)61492-0

11. European Centre for Disease Prevention and Control (ECDC). Rapid Risk Assessment: Severe respiratory disease associated with Middle East respiratory syndrome coronavirus (MERS-CoV). Stockholm: ECDC; 17 May 2013. Available from: http://www.ecdc.europa.eu/en/publications/publications/risk-assessment-middle-east-respiratory-syndrome-coronavirus-mers-cov-17-may-2013.pdf

12. Farrington CP, Kanaan MN, Gay NJ. Branching process models for surveillance of infectious diseases controlled by mass vaccination. Biostatistics. 2003;4(2):279-95. http://dx.doi.org/10.1093/biostatistics/4.2.279

13. De Serres G, Gay NJ, Farrington CP. Epidemiology of transmissible diseases after elimination. Am J Epidemiol. 2000; 151(11): 1039-48. http://dx.doi.org/10.1093/oxfordjournals.aje.a010145

14. Ferguson NM, Fraser C, Donnelly CA, Ghani AC, Anderson RM. Public health risk from the avian H5N1 influenza epidemic. Science. 2004;304(5673): 968-9. http://dx.doi.org/10.1126/science.1096898

15. ProMED-mail. MERS-COV - EASTERN MEDITERRANEAN (31): Jordan, retrospective case identification, WHO, request for information. Archive Number: 20130617.1777989. 17 Jun 2013. Available from: http://www.promedmail.org/direct.php?id=20130617.1777989

16. Fraser C, Donnelly CA, Cauchemez S, Hanage WP, Van Kerkhove MD, Hollingsworth TD, et al. Pandemic potential of a strain of influenza A(H1N1): early findings. Science. 2009; 324(5934): 1557-61. http://dx.doi.org/10.1126/science.1176062

17. Balcan D, Hu H, Gonçalves B, Bajardi P, Poletto C, Ramasco JJ, et al. Seasonal transmission potential and activity peaks of the new influenza A(H1N1): a Monte Carlo likelihood analysis based on human mobility. BMC Med. 2009;7: 45. http://dx.doi.org/10.1186/1741-7015-7-45

18. Colizza V, Barrat A, Barthelemy M, Valleron A-J, Vespignani A. Modeling the Worldwide Spread of Pandemic Influenza: Baseline Case and Containment Interventions, PLoS Med. 2007;4(1):e13. http://dx.doi.org/10.1371/journal.pmed.0040013

19. Balcan D, Colizza V, Gonçalves B, Hu H, Ramasco JJ, Vespignani A. Multiscale mobility networks and the spatial spreading of infectious diseases, Proc Natl Acad Sci USA. 2009;106(51):21484-9. http://dx.doi.org/10.1073/pnas.0906910106

20. Anderson RM, May RM. Infectious Diseases of Humans: Dynamics and Control. Oxford: Oxford University Press; 1992.

21. Center for Disease Control and Prevention (CDC). A Novel Coronavirus Called "MERS-CoV" in the Arabian Peninsula, updated July 26, 2013. Atlanta: CDC Available from: http://wwwnc.cdc.gov/travel/notices/watch/coronavirus-saudi-arabia-qatar#footnote

22. Venzon DJ, Moolgavkar SH A method for computing profile-likelihood based confidence intervals. Applied Statistics. 1988; 37: 87-94. http://dx.doi.org/10.2307/2347496

23. International Civil Aviation Organization and Civil Aviation Statistics of the World, statistics. Washington, DC: The World Bank. Available from: http://data.worldbank.org/indicator/IS.AIR.PSGR

24. World Health Organization (WHO), Global Alert and Response (GAR), MERS-CoV update – as of 26 June 2013. Geneva: WHO. Available from: http://www.who.int/csr/don/2013_06_26/en/index.html

25. World Health Organization (WHO), Global Alert and Response (GAR), Middle East respiratory syndrome coronavirus (MERS-CoV) - update 7 September 2013. Available from: http://www.who.int/csr/don/2013_09_07/en/

26. Holmes KV. Structural biology. Adaptation of SARS coronavirus to humans. Science. 2005; 309(5742): 1822–23. http://dx.doi.org/10.1126/science.1118817

27. Khan K, Sears J, Hu VW, Brownstein JS, Hay S, Kossowsky D, et al. Potential for the International Spread of Middle East Respiratory Syndrome in Association with Mass Gatherings in Saudi Arabia. PLOS Curr. 2013;5.

28. Gautret P, Charrel R, Belhouchat K, Drali T, Benkouiten S, Nougairede A, et al. Lack of nasal carriage of novel corona virus (HCoV-EMC) in French Hajj pilgrims returning from the Hajj 2012, despite a high rate of respiratory symptoms. Clin Microbiol Infect. 2013; 19(7):E315-7. http://dx.doi.org/10.1111/1469-0691.12174

29. World Health Organization (WHO), International travel and Health, World-travel advice on MERS-CoV for pilgrimages, 25 July 2013. Geneva: WHO. Available from: http://www.who.int/ith/updates/20130725/en/index.html

30. Hufnagel L, Brockmann D, Geisel T. Forecast and control of epidemics in a globalized world. Proc Natl Acad Sci USA. 2004; 101(42):15124. http://dx.doi.org/10.1073/pnas.0308344101

31. Colizza V, Barrat A, Barthelemy M, Vespignani A. Predictability and epidemic pathways in global outbreaks of infectious diseases: the SARS case study. BMC Med. 2007;5: 34. http://dx.doi.org/10.1186/1741-7015-5-34

32. Khan K, Arino J, Hu W, Raposo P, Sears J, Calderon F, et al. Spread of a Novel Influenza A (H1N1) Virus via Global Airline Transportation. N Engl J Med. 2009; 361(2): 212-4. http://dx.doi.org/10.1056/NEJMc0904559

33. Tizzoni M, Bajardi P, Poletto C, Ramasco JJ, Balcan D, Gonçalves B, et al. Real-time numerical forecast of global epidemic spreading: case study of 2009 A/H1N1pdm. BMC Med. 2012;10:165. http://dx.doi.org/10.1186/1741-7015-10-165