Forecasting versus projection models in epidemiology: The case of the SARS epidemics

Eduardo Massad [a,b,*], Marcelo N. Burattini [a], Luis F. Lopez [a], Francisco A.B. Coutinho [a]

[a] *School of Medicine, The University of São Paulo, Av. Dr. Arnaldo 455, São Paulo, 01246-903, SP, Brazil*
[b] *Department of Infectious and Tropical Diseases, London School of Hygiene & Tropical Medicine, London, UK*

Summary   In this work we propose a simple mathematical model for the analysis of the impact of control measures against an emerging infection, namely, the severe acute respiratory syndrome (SARS). The model provides a testable hypothesis by considering a dynamical equation for the contact parameter, which drops exponentially with time, simulating control measures. We discuss the role of modelling in public health and we analyse the distinction between forecasting and projection models as assessing tools for the estimation of the impact of intervention strategies. The model is applied to the communities of Hong Kong and Toronto (Canada) and it mimics those epidemics with fairly good accuracy. The estimated values for the basic reproduction number, R_0, were 1.2 for Hong Kong and 1.32 for Toronto (Canada). The model projects that, in the absence of control, the final number of cases would be 320,000 in Hong Kong and 36,900 in Toronto (Canada). In contrast, with control measures, which reduce the contact rate to about 25% of its initial value, the expected final number of cases is reduced to 1778 in Hong Kong and 226 in Toronto (Canada). Although SARS can be a devastating infection, early recognition, prompt isolation, and appropriate precaution measures, can be very effective to limit its spread.
© 2005 Elsevier Ltd. All rights reserved.

Introduction

Public health is the technological arm of epidemiology and its actions are based essentially on prevention and intervention. To prevent and to intervene, however, one must predict. To predict the natural course of a system in the absence of intervention and to predict what is going to happen to such a system after the proposed intervention. Prediction, on its side, is strongly dependent on the scientific foundations of the subject in question. In this sense, mathematical models have provided a precise template against which new observations or theories can be tested as understanding of the phenomenon develops [1].

* Corresponding author. Tel.: +55 11 30817717; fax: +55 11 30667382.
  E-mail address: edmassad@usp.br (E. Massad).

Three major aims of mathematical models in epidemiology can be identified: the first centres on the need for scientific understanding and precision in the expression of current theories and concepts; a second aim, linked to the first, is the role of theory in identifying areas in which better epidemiological data is required to refine prediction and improve understanding; and the third, and in many instances, the most difficult objective is that of prediction [1]. In addition to these three aims of modelling we propose a fourth objective: the generation of testable hypotheses by providing a theoretical framework on which plausible scenarios can be simulated in a computer environment (*in silico* experiments).

Prediction in general science can be divided into two components: forecasting and projections [2]. A forecast is an attempt to predict what *will* happen. A projection is an attempt to describe what *would* happen, given certain hypotheses [3]. Among the tools available to the modern epidemiologists for both forecasting and projection are the mathematical (or dynamical) models, which, when well structured, can provide predictive capacity to the public health professional, helping in the design, and assessment of the impact, of control strategies [4—7]. For instance, by projecting what would happen with a given population if individuals were not vaccinated, it is possible to quantify the relative impact of a specific vaccination program.

The aim of this work is to provide a projection of what would have happened with the course of severe acute respiratory syndrome (SARS) epidemic if the universal procedures to reduce contact were not implemented in the affected areas. We do this by the application of a dynamical system approach, as described below.

The SARS epidemics: a case in point

SARS is a recently discovered infectious disease with high potential for transmission [8], transmitted by droplet and direct contact and caused by a new strain of corona virus [9]. On 5 July 2003, WHO announced that the last known chain of human-to-human transmission of the SARS corona virus had been broken. A cumulative number of 8422 cases have been reported worldwide to the WHO [10], with 908 deaths, as of August 2003.

In the end of 2002, reports from China suggested that a new, highly contagious, and very severe atypical pneumonia of unknown cause was occurring in the Guangdong province. As it reached southeastern Asian countries, the condition appeared to be particularly prevalent among health care workers and their household members. In response to that threat, on 13 March 2003, WHO issued a global alert, for the first time on more than a decade, and instituted worldwide surveillance. On March 27, scientists in the WHO laboratory network reported major progress in the identification of the causative agent, a new member of the corona virus family.

By that time, SARS has already become a global health hazard, and its high infectivity was alarming. Early recognition, prompt isolation, and appropriate precaution measures were considered to be key factors in combating this infection [11]. However, while much has been learnt about SARS since it was brought to international attention in March 2003, there remain many unanswered questions about where it came from, how it spreads, and the effectiveness of public health and other measures employed to control the disease [12].

This paper is an attempt to answer the last question. For this, we propose a dynamical model to assess the expected burden of the epidemic in the absence of control measures and the impact of adopting well-known precautions methods, like the set of measures included in the air-borne and contact precautions recommended by CDC [9], on the epidemic course.

The model and results

Some affected communities, like Hong Kong and Toronto (Canada), have a detailed record of the epidemics, allowing a deeper analysis of its dynamics. We designed a simple mathematical model to describe the spread of SARS and to predict the impact of control measures. We calculated the basic reproduction number, R_0, for those communities and simulated the temporal evolution of the epidemics with and without control measures.

The model is deterministic of the susceptible—infected—recovered (SIR) [13] type, assuming homogeneously mixing contacts and considering the known parameters (duration of the infectious period and mortality rate of the infection). The model is described by the following set of differential equations, representing the temporal evolution of the number of individuals in each of the possible states, susceptibles $X(t)$, infectious, $Y(t)$ and recovered, $Z(t)$:

$$dX(t)/dt = -\beta X(t)Y(t)/N(t) - \mu X(t) + \mu[Y(t) + Z(t)],$$
$$dY(t)/dt = \beta X(t)Y(t)/N(t) - [\mu + \gamma + \alpha]Y(t),$$
$$dZ(t)/dt = \gamma Y(t) - \mu Z(t),$$

$$(1)$$

where β is the contact rate, γ is the recovery rate, μ is the natural mortality rate and α is the infection mortality rate. The cumulative number of cases, $C(t)$, was calculated according to

$$C(t) = \int_0^t [\beta X(s)Y(s)/N(s)] \, ds. \qquad (2)$$

The contact rate parameter was estimated from the exponential initial phase of the epidemics [13−16]. As we mentioned in a previous work [17], R_0 can be estimated from the initial exponential growing phase of the number of cases. For this, we begin by fitting an exponential curve to the initial growing phase of the number of human cases of SARS, $Y(t)$

$$Y(t) = Y_0 \exp[(R_0 - 1)t]. \qquad (3)$$

Having R_0 it is possible to estimate the contact rate, β, the most difficult parameter to be obtained from epidemiological field research, through the relation

$$\beta = R_0(\mu + \gamma + \alpha). \qquad (4)$$

We then simulated the model numerically under two hypothetical scenarios, the natural course of the infection (the contact parameter is kept constant in time) and under control measures (the contact parameter drops exponentially with time).

In Fig. 1 we show the simulation for the Hong Kong community. The parameters applied were $\mu = 0.00004$ days^{-1}; $\gamma = 0.1$ days^{-1}; and $\alpha = 0.08$

days^{-1}, adapted from the data described by the WHO [18]. The initial value of β was calculated according to Eq. (4) and we used the relation

$$\beta = 2.16 \times 10^{-1} \text{ for } t \leqslant 15 \text{ days, and}$$
$$\beta(t) = (2.16 \times 10^{-1}) \exp(-\kappa t) \text{ for } t > 15 \text{ days,}$$

$$(5)$$

with $\kappa = 0.0158$ days^{-1}, for simulating control. That is, for the first 15 days β was assumed as constant and equal to 0.216 potentially infective contacts per day, dropping exponentially with rate $\kappa = 0.0158$ days^{-1} thereafter. In order to compare the actual curve of cases with the theoretical projection of the number of cases without any control we kept the rate κ as equal to zero for the entire epidemic period, as shown by the continuous line in Fig. 1.

In Fig. 2 we show the simulation for Toronto (Canada). The Canadian case has some peculiarities worth commenting. Besides being the only western country with a significant autochthonous SARS epidemic, after about 65 days the beginning of the first outbreak, when the number of new cases was almost zeroed, a second outbreak began, as can be easily noted by the sudden rise in the number of new cases occurring between days 65 and 80 (see Fig. 2 below). The parameters used were

$$\beta = \beta(t) = (2.9 \times 10^{-1}) \exp(-\kappa t), \qquad (6)$$

with $\kappa = 0.0145$ days^{-1}, simulating control; $\mu = 0.00004$ days^{-1}; $\gamma = 0.1$ days^{-1}; and $\alpha = 0.12$ days^{-1}. The second outbreak was simulated by assuming the number of the last cases before the outbreak as initial condition and applying the same equation for $\beta(t)$, with $\beta(0) = 2.9 \times 10^{-1}$ but with $\kappa = 0.45$ days^{-1}. We therefore assumed a very quick
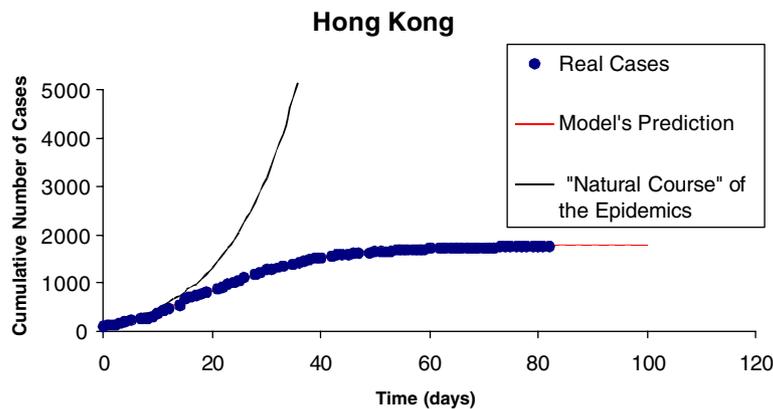

Hong Kong

Figure 1 Results of the simulation for the Hong Kong community. The parameters applied were: $\beta = \beta(t) = (2.6 \times 10^{-7}) \exp(-\kappa t)$, with $\kappa = 0.0158$ days^{-1}, simulating control; $\mu = 0.00004$ days^{-1}; $\gamma = 0.1$ days^{-1}; and $\alpha = 0.08$ days^{-1}. The same community was simulated with the line expressing the ''Natural Course'' of the epidemics calculated with $\kappa = 0$, simulating the absence of control.
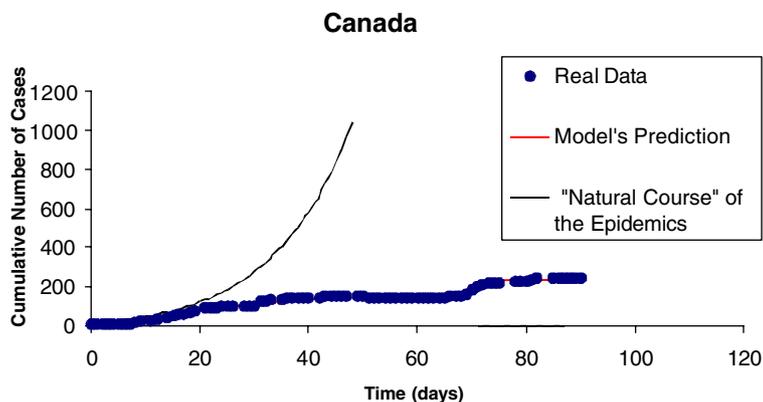
Canada

Figure 2 Results of the simulation for Toronto (Canada). The parameters applied were: $\beta = \beta(t) = (2.9 \times 10^{-7}) \exp(-\kappa t)$, with $\kappa = 0.0145$ days^{-1}, simulating control; $\mu = 0.00004$ days^{-1}; $\gamma = 0.1$ days^{-1}; and $\alpha = 0.12$ days^{-1}. The second outbreak was simulated by assuming the number of the last cases before the outbreak as initial condition and applying the same equation for $\beta(t)$ but with $\kappa = 0.45$ days^{-1}. We therefore assumed a very quick response of the Canadian health authorities to this second outbreak. The same community was simulated with the line expressing the ''Natural Course'' of the epidemics calculated with $\kappa = 0$, simulating the absence of control.

response of the Canadian health authorities to this second outbreak.

The calculated value for R_0 for Hong Kong was 1.20 and for Toronto (Canada) was 1.32. In Fig. 3 we show the simulation of the reproduction number of both epidemics.

At time equals zero we have the basic reproduction number, R_0, and the time evolution of the number of secondary cases. We note that this value drops below the threshold due to the exponential reduction in the contact parameter β, as explained above. It can be noted also the recrudescence of the Canadian epidemic, when the value of β re-

turned to its basal value. When the reproduced number crosses the threshold line ($R = 1$), the number of new cases reaches its maximum, dropping thereafter.

The model mimics real data with good accuracy for both communities when considering adoption of control measures. The model's prediction demonstrated an epidemic that is, by far, milder than expected without control measures. The model projects that, in the absence of control, the final number of cases would be 320,000 in Hong Kong and 36,900 in Toronto (Canada). By contrast, with control measures, which reduce the contact rate

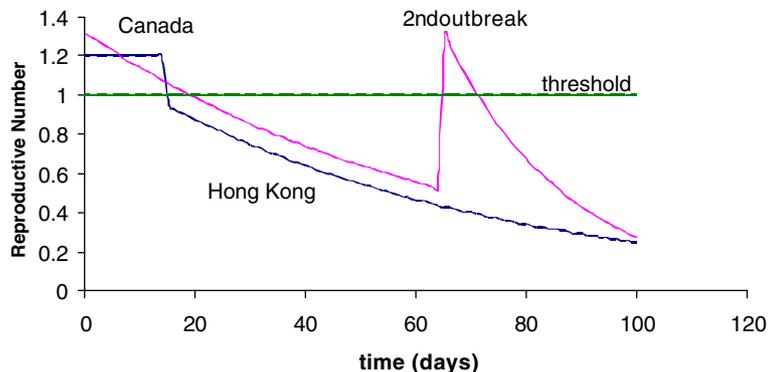**Time evolution of SARS Reproduction Number**

Figure 3 The reproduction number, $R = R_0[X(t)/N(t)]$ expressing the number of secondary cases for both communities. We note that, at time equals zero we have the value of the basic reproduction number, R_0, and this number of secondary cases drops with the reduction in the number of available susceptibles and, in our case, with the exponential reduction in the contact parameter β.

to about 25% of its initial value, the expected final number of cases is reduced to 1778 in Hong Kong and 226 in Toronto (Canada). In fact, the stability level predicted by the model was indeed attained in both Hong Kong and Toronto (Canada) by the end of the outbreaks.

Note that the model's performance should be taken with some caution, since there are some constraints that could influence our results. The first one is the assumption of homogeneously mixing pattern of transmission. In fact, this is a simplification that does not take into account the fact that SARS occurred mainly as focal outbreaks, with a concentration of cases in small communities somewhat related. This simplification may overestimate the projections of the natural course of the epidemics. However, the qualitative results are not affected, that is, by introducing in the model the parameter κ in Eq. (5), which reduced the contact rate, we were able to reproduce the change in the natural impetus of the epidemics. The second constraint of the model is the deterministic nature of the model, that is, we considered only the average values of each variable and parameter (the mean field approach). However, the size of the epidemics and the affected populations justify the deterministic approximation.

Finally, as our objective was not to forecast the real size of the epidemics but rather to analyse the possible impact of implementing control measures, the above-mentioned constraints do not compromise the final conclusions.

Discussion

This work proposes a theoretical framework of the SARS epidemic dynamics on which a hypothesis, namely, the spread of the infection can be checked by simple, universal procedures that reduces contacts, was tested. The test consisted in reproducing the course of the epidemic along its temporal evolution by mimicking the contact reduction in the parameter β, as compared with the absence of intervention. The model reproduces the temporal evolution of the epidemics with good accuracy and suggests what would have happen if the control measures were not implemented. Therefore, our model is not a forecasting model but rather a projection one. This distinction is an important one, in particular concerning the concept of validation. Forecasting models can be tested by comparing their predictions with actual data. Projection models should be seen in a counterfactual way, that is, their projections cannot be tested because they will never happen.

The main purpose of projection models is, therefore, to provide support by plausible hypotheses on the efficacy of intervention policies. This is the role of the present model, to provide support to the hypothesis that simple control procedures interrupted the transmission chain of the SARS virus in the areas analysed.

In spite of the fact that SARS is indeed a very contagious disease, public health authorities have succeeded in restricting its spread, by isolating patients and using measures to block exposure to coughed-up droplets, such as face-masks, which also avoid indirect contact by limiting hand—nose—hand—environment contamination. Those are the same well known measures applied to limit transmission of other directly or air-borne transmitted infections such as tuberculosis or influenza.

The novel aspect of our approach is the proposition of a method for mimicking control measures, represented by the function that describes the reduction in the contact probability with time. The model reproduced the epidemic in Hong Kong and the Toronto case, where the epidemic recrudesced. No explanation for this second outbreak has been proposed. We suggest that it was probably due to a relaxing of public health measures (β immediately returned to β_0), which were promptly corrected with a greater rigour than that applied during the first outbreak (κ increased 30-fold). This phenomenon can be seen in the curves of Fig. 3, which display the time evolution of the number of secondary cases. As mentioned above, this number starts with the basic reproduction number, R_0, dropping according with the reduction in the proportion of susceptibles and, in our case, also with the reduction in the contact parameter β.

Another aspect of our model worth commenting is the absence of a latency compartment. Although variable, short incubation periods of less than a week has been considered to be dominant in SARS [19,20]. In addition, most of the incubating cases evolved to clinical cases and were probably infectives before clinical recognition, what reinforces our SIR structure for the model. In fact, recent publications [21,22] consider that sub clinical SARS is not an important feature of the disease.

But perhaps the most important contribution of this work is that our approach allows the projections of what would have happened if control measures had not been implemented with the appropriate speed and efficacy. As seen in Figs. 1 and 2, the SARS epidemics had, indeed, the potential of reaching a huge number of cases, spreading havoc among the affected populations. However, super spreading events are suspected to have a pivotal role in the global spread of SARS [21—23].

In fact, it is suspected to have played major roles in epidemics of Hong Kong and Toronto (Canada). Therefore, it is possible that our projections of the size of those epidemics can be overestimated since we assumed homogenously mixing and our calculations could not be generalized for the entire population of those communities due to the heterogeneity in spreading potential of the affected clusters of individuals.

Although SARS can be a devastating infection, our model shows that simple measures can be very effective to limit its spread. Simple dynamic models can indeed be very helpful in providing insights to public health authorities, in particular in situations where very little is known about the course of emerging or re-emerging epidemics.

Acknowledgements

References

[1] Anderson RM. Epidemiological models and predictions. Trop Geogr Med 1988;40(3):S30—9.

[2] Keyfitz N. On future population. J Am Stat Assoc 1972;67:347—63.

[3] Caswell H. Matrix population models. Sunderland, MA: Sinauer Associates, Inc. Pub.; 1989.

[4] Amaku M, Coutinho FAB, Azevedo RS, Burattini MN, Lopez E, Massad E. Vaccination against rubella: analysis of the temporal evolution of the age-dependent force of infection and the effects of different contact patterns. Phys Rev E 2003;67(051907):1—11.

[5] Burattini MN, Coutinho FAB, Lopez LF, Massad E. Modelling the dynamics of Leishmaniasis considering human, animal host and vector populations. J Biol Syst 1998;6(4):337—56.

[6] Massad E, Azevedo-Neto RS, Burattini MN, Zanetta DMT, Coutinho FAB, Yang HM et al. Assessing the efficacy of a mixed vaccination strategy against rubella in São Paulo, Brazil. Int J Epidemiol 1995;24(4):842—50.

[7] Burattini MN, Massad E, Coutinho FAB. Dynamics of malarial transmission based on serological data. Epidemiol Infect 1993;111(3):503—23.

[8] WER. Severe acute respiratory syndrome (SARS). Wkly Epidemiol Rec 2003;78:1—3.

[9] CDC. Severe acute respiratory syndrome: what everyone should know about SARS. Centers for Disease Control and Prevention. Available from: http://www.cdc.gov/ncidod/sars/basics.htm April 28, 2003.

[10] WHO. Cumulative number of reported probable cases of severe acute respiratory syndrome (SARS). World Health Organization. Available from: http://www.who.int/csr/sarscountry//en/ June 9, 2003.

[11] Lee N, Hui D, Wu A, Chan P, Cameron P, Joynt GM et al. A major outbreak of severe acute respiratory syndrome in Hong Kong. New Engl J Med 2003;348(20): 1986—94.

[12] WHO. Severe acute respiratory syndrome (SARS) — Report by the Secretariat. World Health Organization Executive Board, EB113/33. Available from: http://www.who.int/csr/don/2003_07_04/en/print.html March 2004.

[13] Anderson RM, May RM. Infectious diseases of humans. Oxford: Oxford University Press; 1991.

[14] Marques CA, Forattini OP, Massad E. The basic reproduction number for Dengue fever in São Paulo state, Brazil: 1990—1991 epidemics. Trans Roy Soc Trop Med Hyg 1994;88(1):58—9.

[15] Massad E, Burattini MN, Ortega NRS. Fuzzy logic and measles vaccination: designing a control strategy. Int J Epidemiol 1999;28(3):550—7.

[16] Massad E, Coutinho FAB, Burattini MN, Lopez LF. The risk of Yellow fever in a Dengue infested area. Trans Roy Soc Trop Med Hyg 2001;95(3):370—4.

[17] Massad E, Burattini MN, Coutinho FAB, Lopez LF. Dengue and the risk of urban Yellow fever reintroduction in São Paulo, Brazil. Braz J Med Biol Res 2003;37(4):477—84.

[18] WHO. Severe acute respiratory syndrome (SARS) — Status of the outbreak and lessons for the immediate future. World Health Organization. Communicable disease surveillance and response, 20 May 2003.

[19] Meltzer MI. Multiple contact dates and SARS incubation periods. Emerg Infect Dis 2004;10(2):207—9.

[20] Chang-Yeung M, Xu RH. SARS: Epidemiology. Respirology 2003;8(Suppl S NOV):S9—S14.

[21] Chow PKH, Ooi EE, Tan HK, Sil BK, Teo M, Ng T et al. Healthcare worker seroconversion in SARS outbreak. Emerg Infect Dis 2004;10(2):249—50.

[22] Hsieh YH, Chen CWS, Hsu SB. SARS outbreak, Taiwan, 2003. Emerg Infect Dis 2004;10(2):201—6.

[23] Shen Z, Ning F, Zhou WG, He X, Lin CY, Chin DP et al. Superspreading SARS events, Beijing, 2003. Emerg Infect Dis 2004;10(2):256—60.