Proceedings of the 5th World Congress on Intelligent Control and Automation, June 15-19, 2004, Hangzhou, P.R. China

第五届全球智能控制与自动化大会
2004年6月15 - 19日, 中国杭州

Research on Mathematical Transmission Model of SARS Based on Molecular Kinematics *

Li Jinping Yin Jianqin Cao Aizeng Dong Jiwen Yang Bo
School of Information Science and Engineering, Jinan University, Jinan, Shandong Province, PR China, 250022
lijp_ise@ujn.edu.cn, lijp2001@vip.sina.com

Abstract: **Based on molecular kinematics and other related jobs, an explicit analytical mathematical model about severe acute respiratory syndrome (SARS) and the similar epidemics is put forward. New factors are considered such as resistibility of individuals and system of SARS checkpoints and health registration. Simulation based on molecular kinematics shows the analytical model is able to describe the dissemination of SARS correctly before and after the applications of various control measures, and therefore is able to predict the future epidemics. Application of the model to Beijing's epidemics presents satisfactory results. Principle of superposition is pointed out in determining the value of *R* in open model. Another useful conclusion is that the system of checkpoints and health registration plays critically important role. Drawbacks of the model are also discussed.**

Index Terms: **SARS, Analytical Mathematical Model of Epidemics, Molecular Kinematics, Simulation.**

I INTRODUCTION

Severe acute respiratory syndrome (SARS) is a recently reported illness that has spread widely since the first case was found in Guangdong on 16 Nov 2002. Large-scale outbreaks took place in Guangdong during 3 and 14 Feb 2003, where 305 cases were reported on 11 Feb, subsequent outbreaks were also reported in Beijing, Shanxi province, Inner Mongolia *et al.* Meanwhile, SARS spread rapidly over more than 31 countries around the world via air travel[1].

Great progress has been made to date about SARS[2]. It was found that a novel coronavirus results in SARS (*i.e.*, SARS-Cov), the full sequence of RNA of the virus was then reported on 13 Apr 2003[3,4]. Various effective and stringent control measures were meanwhile taken in place, which greatly prevented the further dissemination and significantly reduced the basic reproductive number *R*, defined as the expected number of secondary infectious cases generated by an index case in a susceptible population[5,6]. The control measures include: establishment of system of SARS checkpoints and health registration, regular measurement of body temperature, isolation of SARS symptomatic cases, quarantine and close observation of asymptomatic contacts, regular disinfection of habitation *et al.* Meanwhile, public awareness of health greatly improves the people's routine habits, such as regular ventilation, actively reducing the social activities, taking more regular exercises[7]. As of the end of May, the number of SARS cases declined down to less than 10 cases everyday[7,8].

Practice shows the most effective measures to block SARS are system of checkpoints, isolation of symptomatic cases, quarantine of asymptomatic contacts[5-9]. Of course this does not decline the impact of other measures. It also shows that the earlier the control measures are taken, the more significant the effect of blocking SARS is. To stop an outbreak, *R* must be maintained below 1[5,6]. These valuable experiences are also indicated in recent researches, most of which focus on the biological properties of coronavirus *et al*[10]. Some jobs strove to study the transmission dynamics and the impact of different control measures[5,6,11-14]. Detailed analysis are performed on the estimation of *R*, especially how the various interventions take effect. Various actual factors such as the variation of *R*, super spread events (SSE), the transmission probability by asymptomatic persons *et al* are considered[5,6]. These realistic works do not provide explicit analytical mathematical model of SARS dissemination. On the other hand, explicit analytical mathematical research can not be easily found in our preliminary investigation of epidemics or infectious illness through library, electronic books and web sites, except several researches on model of epidemics, most of which come from university journals[11], therefore great efforts should be paid to the mathematical research of epidemic model.

Researches on epidemic model can be classified into two categories: one is based on the classical work of Anderson and May by construction and analysis of epidemic differential equations[15,16], the other takes Monte Carlo simulation as its basic approach, which is often a stochastic model and based on Markov process[17]. Based on the realistic analysis, another kind of mathematical model is advanced in the present paper, similar to an experienced model.

* Supported by fund of provincial academic young backbones in colleges and universities of Shandong. **Li Jinping**, a professor of school of information science and engineering (ISE) of Jinan University, he received Ph.D of theoretical physics from University of Science and Technology of China (USTC) in 1997, his present interests include artificial intelligence, image processing, pattern recognition, optimization algorithms. **Yin Jianqin**, an instructor of ISE of Jinan University, she received master's degree of control theory in Shandong University in 2002, her present interests are image processing, optimization algorithms. **Cao Aizeng**, an instructor of ISE of Jinan University, he received master's degree in Shandong Normal University in 2003, his present interests involve image processing, algorithm analysis. **Dong Jiwen**, a vice professor of ISE of Jinan University, he received master's degree of computer in Wuhan Industrial University in 1995, his present interests are image processing, computational materials. **Yang Bo**, a professor of ISE of Jinan University, he received master's degree of computer in 1992 in Southeast University, his present interests are computer networks, information security, distributed computation.

Just like thermodynamics and statistical physics, the study of SARS transmission can also have two major choices: one from macroscopic viewpoint, the other from microscopic viewpoint. The former is based on the average analysis on the whole, the model of epidemics can be acquired from the data of the National Ministry of Health or the Beijing's center for disease control and prevention (BJCDC)[7,8], artificial neural networks is a possible choice for analysis. On the other hand, the model can also be acquired by an essential hypothesis that one infectious individual infects the same number of secondary reproductive infectious cases, *i.e.* R, in the absence of control measures and after the application of full scale interventions. In general, $R>1$ in the early stage of epidemics and $R<1$ when the various control measures take effect (see the detailed analysis in *B* of section II). The latter scenario is a similar formulation to references[5,6,11], which begin with the analysis of general social behavior of infectious individuals in a population of a compact district, similar to the molecular kinematics that studies the collective properties of a molecule system by tracing each molecule. Note that only the infectious individuals not being isolated contribute to the dissemination of SARS; the susceptible population do not directly contribute to SARS spread, they have the possibility of being infected when they have intimate contact with infectious cases; the hospitalized and the dead individuals do not cause further transmission; the recovered and the incubative ones are not found to be infectious, and there is no direct evidence of transmission from an asymptomatic person at present. Thus we only consider the infectious and the susceptible individuals, who can be regarded as traced molecules, the contacts between an infectious case and the susceptible are similar to the collision between motional molecule and stationary molecules.

Since the models are not directly obtained by thorough analysis from the data of Beijing, Hong Kong, Singapore, *et al*, the models are simplified ones. However they are consistent with the actual development of SARS in Beijing.

II MODEL

A. Definitions

Serial interval: the time from the onset of symptoms in an index case to the onset of symptoms in a subsequent case infected by the index case.

Basic reproductive number of an infection: defined as the expected number of secondary infectious cases generated by an average infectious case in an entirely susceptible population, denoted as R. R may be much greater than 1 in the early stage of epidemics, and shall be less than 1 when the control measures are taken in place to stop the epidemics.

Index cases: a special introduced infectious patient by whom the infected individuals can be traced and investigated for the transmission of epidemics in a susceptible population.

Super spread events (SSE): a situation in which a single individual has directly infected a large number of other people.

B. Mathematical Model of SARS

Exponential growth in the number of infectious cases is an often adopted scheme in the early stage of epidemics, *i.e.*,

$$y_1=[n_0 R^{\frac{x}{L}}] \tag{1}$$

This can be completely determined by the actual data from a specific city or district or some countries. Note $[x]$ means the integral part of x, y_1 is the total number of infectious cases, x denotes days, n_0 is the initial introduction number of SARS patients, L represents the average interval which is determined by the incubation period. The correct value of R is a key problem, and sophisticated discussion was performed in [6], since the impact of many control measures are reflected in the variation of value of R. In this paper, some words about R will be given in subsection *C*. It is more attractive to determine the corresponding explicit form by theoretical approach.

Suppose an infectious person is introduced in a susceptible population. Then $y_1=1$ for $x=0$ shows there is only 1 infectious case at the beginning, and $y_1=1+R$ for $x=L$ shows there are $1+R$ infectious cases altogether when $x=L$ days. When $x=2L$, $y_1=1+R+R^2$, note that only the newly produced R infectious persons can lead to the further transmission, the initial introduced case will no longer contribute to the further epidemic spread, since it has caused R secondary infectious cases. Similar analysis can be performed for $x=3L$, $4L$, ... When $x=kL$, $y_1=1+R+R^2+\ldots+R^k = (R^{k+1}-1)/(R-1)$. If the number of introduced SARS cases is n_0, the expression becomes $y_1=n_0(R^{k+1}-1)/(R-1)$. Note that $k = x/L$, the final model in the absence of control measures is

$$y_1=[n_0 \frac{R^{\frac{x}{L}+1}-1}{R-1}] \tag{2}$$

This is a generalization of discrete form to continuous form. Apparently, (2) is different from (1), and (2) comes back to (1) when $R\to\infty$. Note the model only holds when y_1 is far less than the number of susceptible population, since the probability of cross-infectious contacts becomes greater with the augment of the infectious cases y.

At the early stage of epidemics, an open model must be considered, because successive introduced infectious patients deteriorate the epidemics. Suppose n_1, n_2,..., n_k are numbers of introduced infectious cases corresponding to certain duration intervals z_1, z_2, ..., z_k. Then the compound model is

$$y_1 = [\sum_{i=0}^{k} n_i \frac{R^{\frac{x-z_{i+1}}{L}+1}-1}{R-1}] \tag{3}$$

where $z_0=0$. Note the term $n_i(R^{\frac{x-z_i}{L}+1}-1)/(R-1)$ contributes to y_1 only when $x\geq z_i$, and its contribution is zero if $x<z_i$. Of course n_i will vary with the epidemic situation.

With the development of SARS, control measures will certainly be taken, such as isolation of symptomatic cases, quarantine and close observation of asymptomatic contacts so that they may be isolated as soon as they show possible signs of the disease, decline in the time from symptom onset until hospital isolation as much as possible, system of checkpoints and health registration, *et al*. The first 3 measures decrease R remarkably down to below 1, the last one is much helpful in

the prevention of disease from further dissemination to places of mild or no epidemics, which leads to the rapid decline in the sequence n_{k+1}, n_{k+2}, …, n_{k+d}, so the impact of input infectious cases become weaker and weaker and can be neglected eventually. Suppose the control measures are taken at $x=z_{k+1}$.

Then (2) and (3) are modified as

$$y_2=[n_{k+1}\frac{1-R^{\frac{x-z_{k+1}}{L}+1}}{1-R}] \tag{4}$$

$$y_2=[\sum_{i=k+1}^{k+d} n_i\frac{1-R^{\frac{x-z_i}{L}+1}}{1-R}] \tag{5}$$

y_2 is the total newly produced SARS patients when $z_i>z_{k+1}$, it does not contain y_1. Like (3), the term $n_i(1-R^{\frac{x-z_i}{L}+1})/(1-R)$ contributes to y_2 only when $x\geqslant z_i$, its contribution is 0 if $x<z_i$. These requirements for (3) and (5) are conventions, and hold for the similar situation hereafter.

The determination of n_i needs further discussion for $z_i>z_k$. Before the control measures are taken, the growth of infectious cases is approximately exponential, *i.e.*, when the control measures are to be taken at $x=z_{k+1}$, the total infectious cases are

$$N_0=[\sum_{i=0}^{k} n_i\frac{R^{\frac{z_{k+1}-z_i}{L}+1}-1}{R-1}] \tag{6}$$

Note that only the newly infected individuals and the newly introduced cases contribute to further transmission, *i.e.*, there are $N_1=y_1(x=z_{k+1})-y_1(x=z_{k+1}-L)$ newly infected individuals further to infect other susceptible, then the new SARS patients infected by N_1 SARS patients shall obey the same law as (2), so there should be $N_2=y_1(x=z_{k+1}+L)-y_1(x=z_{k+1})$ infected ones stayed in incubation period who may show the symptom of SARS during L days on average from $z=z_{k+1}$ to $z=z_{k+1}+L$, which is assumed to obey a strict exponential growth,

$$y'=m\bar{R}^{\frac{z-z_{k+1}}{L}} \tag{7}$$

where $m=y_1(x=z_{k+1})-y_1(x=z_{k+1}-1)$, and $\bar{R}$ can be determined by $N_2=\int_{z_{k+1}}^{z_{k+1}+L} m\bar{R}^{\frac{z-z_{k+1}}{L}}dz=\frac{mL(\bar{R}-1)}{\ln\bar{R}}$. Note y' denotes the number of newly found SARS infectious patients everyday from $z=z_{k+1}$ to $z=z_{k+1}+L$. These N_2 individuals shall be taken as the introduced cases in the second stage when the control measures are taken at $x=z_{k+1}$, and the newly found SARS cases will be isolated soon after they show the symptoms.

On this occasion, (4) or (5) is no longer convenient for further research, since (5) is a discrete form of introduction intervals, *i.e.*, the introduction cases n_i corresponds to a discrete z_i. However, (7) provides a continuous introduction interval. Similar to (4), we obtain the following scheme,

$$y_3'=m\bar{R}^{\frac{z-z_{k+1}}{L}}\frac{1-R^{\frac{x-z}{L}+1}}{1-R} \tag{8}$$

Obviously, an integral must be performed on (8),

$$y_3=\int_{z_{k+1}}^{x} m\bar{R}^{\frac{z-z_{k+1}}{L}}\frac{1-R^{\frac{x-z}{L}+1}}{1-R}dz=y_{31}(x),\quad x\in[z_{k+1},z_{k+1}+L] \tag{9.1}$$

What is the effective form when $x>z_{k+1}+L$? Obviously the above integral will no longer hold because of supp $y'=[z_{k+1}, z_{k+1}+L]$. The new form is

$$y_3=\int_{z_{k+1}}^{z_{k+1}+L} m\bar{R}^{\frac{z-z_{k+1}}{L}}\frac{1-R^{\frac{x-z}{L}+1}}{1-R}dz=y_{32}(x),\quad x>z_{k+1}+L \tag{9.2}$$

This is a rough scheme and the impact of introduced cases n_{k+1}, n_{k+2}, …, n_{k+d} are neglected. In fact, they can be considered by employing (5). Since these two factors are independent of each other, and they contribute to the total number of SARS cases together, the total number of SARS cases can be obtained by combining (5) and (9)

$$y_2=[\sum_{i=k+1}^{k+d} n_i\frac{1-R^{\frac{x-z_i}{L}+1}}{1-R}]+y_{31}(x)\ (\text{or } y_{32}(x)) \tag{10}$$

Now we present the total formulation as follows:

$$\begin{cases} y=[\sum_{i=0}^{k} n_i\frac{R^{\frac{x-z_i}{L}+1}-1}{R-1}], & x<z_{k+1} \\ y=N_0+y_2, & x\geqslant z_{k+1} \end{cases} \tag{11}$$

For a close model, (2) and (9) seems to be perfect, since n_0 is the only initial input parameter. For open model, (11) seems to be good. But there exist many input parameters, n_i's and z_i's in both (3) and (5), which may result in inconvenience in practice. As a matter of fact, there exist intrinsic relation between n_i's and z_i's, *i.e.*, n_i's are function of z_i's,

$$n_i=f(z_i) \tag{12}$$

which may be determined essentially by the statistical law of the epidemics. (12) is a discrete form, and can be converted into a continuous form, *i.e.*, $n=f(z)$, which means the input infectious cases per unit time. Note (7) is already such a form. Then the universal form of one term in (3) and (5) are

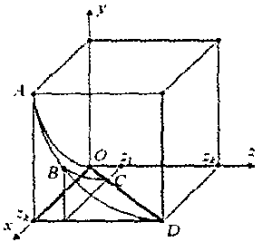

Fig. 1 Effective region (grey) of $y_1(x, z)$. Curves OA and CB represent close models for two situations.

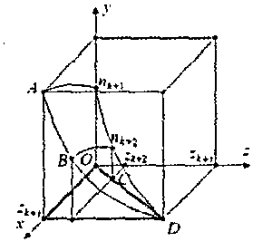

Fig.2 Effective region (gray) of $y_2(x,z)$. Curves $n_{k+1}A$ and $n_{k+2}B$ represent close models for two situations.

$$y_1(x,z)=\frac{f(z)}{R-1}(R^{\frac{x-z}{L}+1}-1),\quad \text{where } z<z_{k+1}, x<z_{k+1} \quad (13)$$

$$y_2(x,z)=\frac{f(z)}{1-R}(1-R^{\frac{x-z}{L}+1}),\quad \text{where } z\geqslant z_{k+1} \quad (14)$$

$x \geqslant z$ is required for both cases. Obviously they are 2 dimensional curving surfaces (see surface surrounded by *ABDCOA* in Fig.1), $x \geqslant z$ always holds in the grey triangular region (see Fig.1-2). We'd like to obtain the total number of infectious cases up to x from (13) and (14). Considering $x \geqslant z$ and the similarity between (13), (14) and probabilistic density distribution, we perform integral on (13) and (14),

$$y_1(x)=\int_0^x \frac{f(z)}{R-1}(R^{\frac{x-z}{L}+1}-1)dz,\quad \text{where } x<z_{k+1} \quad (15)$$

$$y_2(x)=\int_{z_{k+1}}^x \frac{f(z)}{1-R}(1-R^{\frac{x-z}{L}+1})dz,\quad \text{where } x\geqslant z_{k+1} \quad (16)$$

The integrals automatically overcome the previous trouble conventions: terms with $x<z$ have no contribution to $y_1(x)$ or $y_2(x)$. The current question is: whether (15) and (16) can return to their original discrete form? We perform a discretization on (15) and (16), and obtain

$$y_1(x)\doteq\sum_{i=0}\frac{f(z_i)}{R-1}(R^{\frac{x-z_i}{L}+1}-1)\Delta z \rightarrow \sum_{i=0}^{k}\frac{n_i}{R-1}(R^{\frac{x-z_i}{L}+1}-1) \quad (17)$$

$$y_2(x)=\sum_{i=k+1}\frac{f(z_i)}{1-R}(1-R^{\frac{x-z_i}{L}+1})\Delta z \rightarrow \sum_{i=k+1}^{k+d}\frac{n_i}{1-R}(1-R^{\frac{x-z_i}{L}+1}) \quad (18)$$

where "→" represents the introduction of the previously conventions, n_i comes from $f(z_i)\Delta z$ and the intermediate-value theorem. Obviously they are completely the same as (3) and (5), so the answer is yes. The inconvenience is the unknown $n=f(z)$, which shall be determined essentially by the intrinsic statistical law of the epidemics. (15) and (16) are only formalized expression. However, they indeed seem to be perfect, their discrete form (3) and (5) can be used in practice.

We point out that singularity may exist in (11) when $R=1$ because of the denominator $(R-1)$ or $(1-R)$, which can be avoided by setting $R=1+\varepsilon$ or $R=1-\varepsilon$, where ε is a small positive number.

Obviously, this is an explicit analytical mathematical model, and easy for the prediction. Although it seems to be perfect, the correctness must be testified. One approach is to apply it directly to the actual data from Beijing, Hong Kong, *et al*, the other is to perform a simulation based on an actual model. We test the model (11) by both approaches with discretion: firstly, a simulation based on molecular kinematics is carried out, the result is consistent with (11), then by proper settings of parameters, it also complies with the data of Beijing.

C. *Scenario of Simulation Based on Molecular Kinematics*

The outbreak of an infection depends on such necessary conditions as source of infection, route of transmission and herd susceptibility. Other natural and social factors also play important role, *e.g.*, control measures, air temperature and incubation period. For a population, the introduced infectious cases are source of infection. Some animals are also possible source of infection, however, no direct evidences are found [18]. The routes of transmission mainly include airborne and droplet infection, so public places are dangerous. Present researches show all people are susceptible to SARS, so the range of susceptible population is extensive. But it does not mean anyone who has intimate contact with SARS patients will be infected, since different persons have different resistibility. Members in a family or sickroom or hospital with infectious individuals are special susceptible population due to more possibility to be infected.

From the microcosmic viewpoint, the behavior of an individual is well regulated similar to the behavior of a molecule in a gas system where Newton motion laws must be obeyed. However from the macroscopic viewpoint, the social behaviors of all individuals comply with the stochastic process to a great extent. In practice, the social behaviors are collective representations of all individuals in a system.

The infectious and the susceptible individuals can be regarded as traced molecules in a gas system, the contacts between infectious case and the susceptible are similar to the collision between motional molecule and stationary molecules, note that all the contacts between other individuals do not contribute to the epidemics dissemination. Incubation period corresponds to the process of a molecule from motionless to motional, the longer the incubation period, the more difficult for the molecule to move.

Before taking control measures, one introduced infectious individual may infect several persons, usually $R \geqslant 1$, which means the introduced motional molecule has more energy and can make more other motionless molecules to move through collision. Different motionless molecules have different inertia, and have different velocity when collided, which correspond to different resistibility of different susceptible individuals. The SSE correspond to an introduced molecule with enormous energy, which only takes place at the early stage of infection and will not determine the long term dissemination usually, we restrict such molecule to only 1 in the simulation. With more and more introduced infectious patients without control measures, the number of infected ones will grow rapidly, maybe shows an exponential law. This is the ascending stage.

After taking control measures, a platform stage of infection comes, when the infected patients result in few newly infected individuals, *i.e.*, $R<1$, which means the introduced motional molecule has less energy and can not cause more motionless molecules to move. Since there are already many infected individuals, and especially many individuals stayed in incubation period are not quarantined or isolated, the number of SARS cases will stay on a higher level for some days. The equivalent situation for gas system is: many molecules are going to move whereas they appear stationary at the beginning, and the motional molecules are to be taken out as soon as they show motional. With more and more stringent control measures in place, people pay more and more attention to personal hygiene and habitual behavior,

reduce many unnecessary social activities, which correspond to the inertia
of molecules become larger. After some days, the number of infected cases will cease to augment and begin to decline.

Now comes the third descending stage. Many newly found infected individuals are immediately sent to hospitals and isolated, the possible ones who have contacts with the known infectious persons will be quarantined so that they may be isolated as soon as they shows possible signs of the disease, therefore the time from the onset of symptom to hospital isolation decline evidently. This is similar to the extraction of motional molecules, which will no longer have ability to collide with other molecules. On the other hand, the stringent system of checkpoints and health registration prevent the further introduction of infectious patients, this corresponds to the cease of more input motional molecules. Eventually, the disease will disappear.

In the above scenario, six parameters are employed:

1. *Matrix of contact degree* $A[i, j]$. $A[i, j]$ represents the intimacy degree of contacts between individuals i and j, and only the contacts between infectious individual and susceptible individuals contribute to the values of $A[i, j]$. $A[i, j]=1$ stands for the most intimate contact, $A[i, j]=0$ means no contact with SARS coronavirus carriers. From the statistical viewpoint, the initial values of $A[i, j]$ comply with a certain distribution, since no more information can be referred to, we assume it complies with a uniform distribution. Suppose i is an infectious patient, j is a susceptible person, the final contact of j is $\max_i\{A[i, j]\}$.

2. *Matrix of contact duration* $T[i, j]$. $T[i, j]$ represents the contact duration, and only the contacts like in $A[i, j]$ contribute to the values of $T[i, j]$. $T[i, j]=1$ when $T[i, j]$ is greater than a given threshold, $T[i, j]=0$ means no contact with SARS cases. We also assume $T[i, j]$ comply with a uniform distribution.

After control measures are taken, the values of $A[i, j]$ and $T[i, j]$ will decline on the whole.

3. *Matrix of resistibility* $R[j]$. $R[j]$ represents the resistibility of everyone in a susceptible population. We assume it complies with a modified Gaussian distribution,

$$p(x)=Ax(1-x)\exp[-\frac{(x-\mu)^2}{2\sigma^2}] \tag{19}$$

where $x\in[0, 1]$ represents the resistibility, $p(x)$ is probability density of distribution and represents the corresponding proportion in the population, μ stands for the average resistibility, σ is the variance. As for the resistibility of a specific individual, an integral shall be performed on (19),

$$P(x)=\int_0^x p(t)dt \tag{20}$$

This is the distribution of function of probability. A random Y of uniform distribution will produce a special distribution $R[j]$ in compliance with $p(x)$ through $R[j]=x=P^{-1}(Y)$.

After the effective control measures are taken, μ may become a little greater and σ a little less.

4. *Incubation period* L. Clinical data shows $L\in[2,14]$. We assume L complies with Poisson distribution,

$$P(L)=\frac{\lambda^L}{L!}e^{-\lambda} \tag{21}$$

where $L\in\{2,3,4,\ldots 14\}$. Note that $E\{L\}=\lambda$, $D\{L\}=\lambda$, so λ is the average incubation period. Employ the method similar to (20), a random variable in accord with (21) will be generated.

5. *Isolation interval*: the time of infectious patient from the onset of symptom to the hospital.

6. *Successive introduced infectious patients*: the sequence $n_1, n_2, \ldots, n_k, n_{k+1}, \ldots, n_{k+d}$. They are determined by the system of checkpoints and health registration, and obey different rule before and after the application of control measures.

Fig. 3 is the flow chart. The brief interpretations are:

①: *contact range* means the average number of specific susceptible individuals whom an infectious patient may infect, which determine the value of R.

②: used for the loop of program in which each SARS patient infect other susceptible individuals.

③: what are conditions for a susceptible person to be infected? The condition in our simulation is:

$$\max_i\{A[i,j]\}>tha \text{ AND } \max_i\{T[i,j]\}>tht \tag{22}$$

where *tha* and *tht* are two thresholds. In Fig.3, judgment ③ is an abbr. form of (22).

④: *thr* is the threshold of an individual, who can not be infected if his resistibility is greater than *thr*.

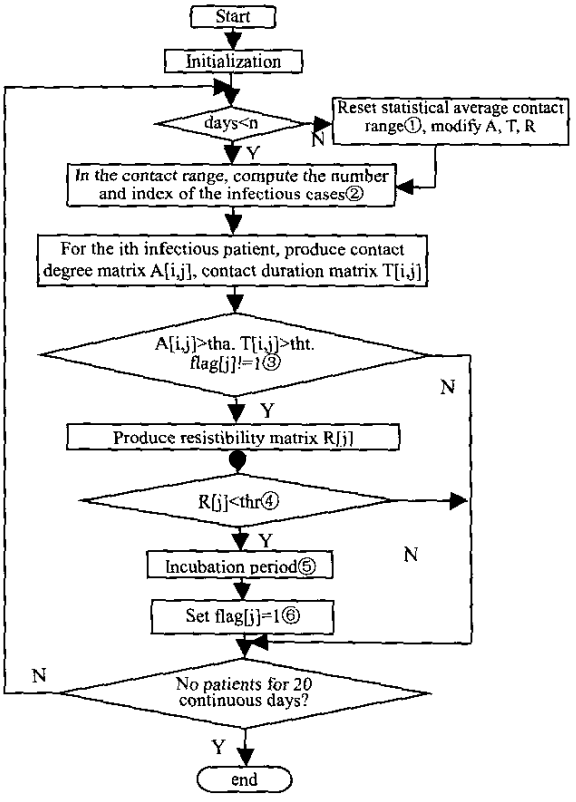

Fig.3 Program flow of simulation based on molecular kinematics

⑤: Incubation period for j is produced according to Gaussian distribution.

⑥: *flag* is an indication. If an individual j is infected, $flag[j]$=1. Otherwise it is set to 0.

In some references[5,6], wide coverage is engaged in the discussion about R. In fact, R is reflected mainly by the number of persons having intimate contact with infectious cases, members in the same family, good friends, doctors or nurses *et al* are susceptible individuals. When various stringent control measures are taken, R will decline significantly.

III Simulation and Discussion

In order to make the simulation more convincible, we perform an average on both the close model and the open model to avoid the accidental coincidence to some extent, *i.e.*, we execute the simulation many times and average the results of each simulation, then the average will be steady and can reflect the essence of the model.

In the simulation, the control measures are taken 45 days later since the first introduced index case. The simulation is performed from 3 aspects: the consistency of simulation with the mathematical model; the consistency of simulation with actual data from Beijing; the extreme situations of simulation if special parameters are used. The first two illustrate the correctness of the present model, the last one will throw light on the impact of control measures, esp. the system of checkpoints and isolation interval. Simulation indicates that it is a disaster without or with much slow and loose control measures, and the epidemics lasts more days when bad system of checkpoints is employed.

How to determine the value of R in mathematical model? Note that model of infection spread obeys the **principle of superposition**: an open model is equivalent to the sum of a series of subsequent close models, a close model induced by n introduced SARS cases is equivalent to another close model induced by one single introduced SARS case by translating the origin to the new position where there are n newly infected patients. The prerequisite is the introduction of SARS cases is independent and the number of SARS cases is far less than the total population. By employing the principle of superposition, R in open model can be easily determined by analyzing the close model, *i.e.*, the value of R in open model is identical with that of close model in the absence of control measures, so does it after the application of control measures. Therefore R shall be determined in close model. First perform the simulation of close model based on molecular kinematics, obtain the total number of SARS cases just before the application of control measures, then calculate R by employing (2). How to get the value of R after the institution of control measures? We adopt a simple and convenient scheme: taking the control measures at the beginning of simulation by providing an initial enormous introduction, *e.g.*, the initial number of introduction cases is set to 500 or 1000, then by employing (4) we can obtain the value of R.

The parameters take the following values:

The scale of population is 1,000,000. Thresholds values: tha=0.8, tht=0.6, thr=0.8, they remained unchanged before and after the institution of control measures. Contact range and isolation interval are set to 7-10 individuals and 10-14 days respectively before the control measures are taken in place, 4-5 individuals and 2-4 days respectively after the control measures underway. For close model, there is only one single introduced SARS case, so $n_1, n_2, \ldots, n_k, n_{k+1}, \ldots, n_{k+d}$=0; for open model, the infection is brought about by the initial introduction of SARS case and the successive introduction of SARS cases, we assume there are 4-6 introduced cases every 3 days in the absence of control measures, and there are 2-3 introduced cases every 6 days and no introduced cases 30-40 days after the institution of control measures.

As discussion in section 2.3, matrix of resistibility and incubation period comply with (19) and (21) respectively, the setting of parameters are: λ=7, μ=0.8, σ=0.3. Then random variables x's can be generated that are in accordance with (19) and (21). In mathematical model, L=7 days.

Fig.4 illustrates the average curve (10 times) of the close model in comparison with the corresponding curve of mathematical model. In the simulation, there is one single introduced infectious SARS case. From Fig.4, we find the average curve is consistent with the mathematical model except the peak around 50 days due to the sharp transition of R in the mathematical model.

Fig.5 presents a special curve corresponding to the much stringent control measures, R=0.001 in mathematical model. It is easy to see the epidemics immediately cease to spread when very stringent control measures are taken in place. The loose control measures result in long term infection (see Fig.4).

Fig.6 shows the average results (10 times) of open model comparing with the corresponding curve of mathematical model, (**C**) and (**D**) provide curves when the control measures are taken 8 days ahead of time. The difference between close and open model is much evident: the epidemics of open model lasts more days and results in much more infected cases when control measures are taken on 45th day, and the curves have longer tails. On the other hand, much less SARS cases are reported when control measures are taken 7 days ahead of time. So the system of checkpoints is important and effective control measures shall be taken in place as early as possible.

Fig.7 presents the comparison of average curves (5 times of total number of SARS cases with 2400-2600) of open model with actual data from Beijing. The scale of abscissa and ordinate has been adjusted for the reasonable comparison. We find favourable consistency of simulation with the actual data. To our delight, we find it useful in accounting for the epidemic situation of Beijing before 21 April, since we can only acquire the actual data of SARS from Beijing after 21 April. It is easy to see that Beijing's infection may begin at the beginning of March (7 March), and there are about 30 cases altogether at the beginning of April. Data in Table 1 are acquired from simulation; we find there may be several hundred SARS cases before 20 April, and 128 cases on 12 April, 27 cases on 2 April. These data of simulation are

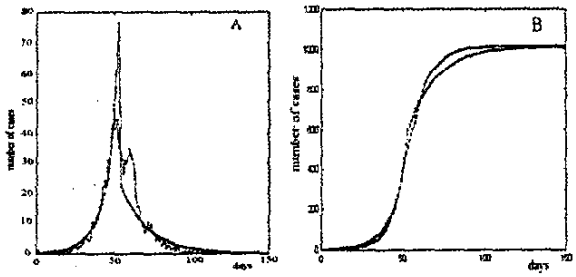

Fig.4 Comparison of close model (blue) with the corresponding mathematical model (red). (A) Curves of number of newly produced SARS cases everyday, (B) Curves of accumulative number of SARS cases. R=2.36 before the control measures, R=0.63 after the control measures, $\widetilde{R}$ =2.59.

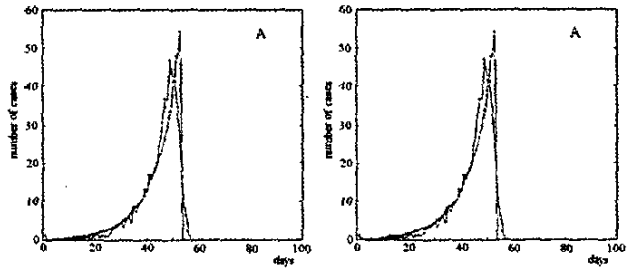

Fig.5 Epidemics of simulation (blue) and mathematical model (red) with very stringent control measures (blue), after the control measures. (A) Curves of number of newly produced SARS cases everyday, (B) Curves of accumulative number of SARS cases. R=2.36 before the control measures, R=0.001 after the control measures.

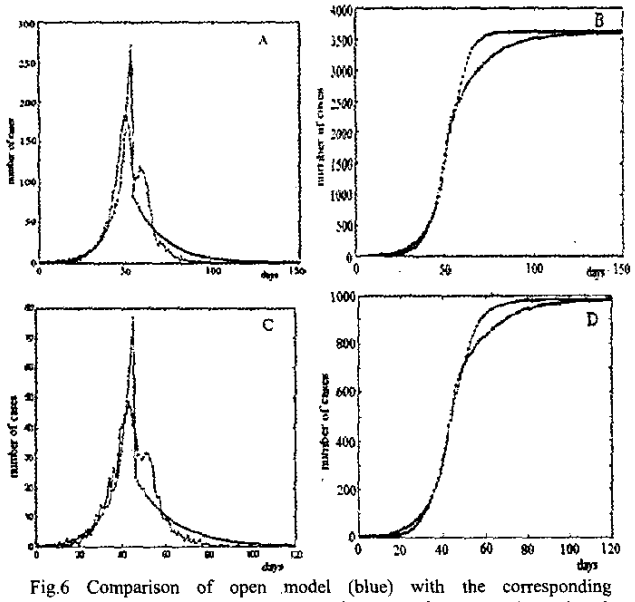

Fig.6 Comparison of open model (blue) with the corresponding mathematical model (red). (A) Curves of number of newly produced SARS cases everyday when control measures are taken on 45th day, $\widetilde{R}$ =2.595. (B) Curves of accumulative number of SARS cases corresponding to (A), (C) Curves of number of newly produced SARS cases everyday when control measures are taken on 37th day, $\widetilde{R}$ =2.62. (D) Curves of accumulative number of SARS cases corresponding to (C). R=2.36 before the control measures, R=0.63 after

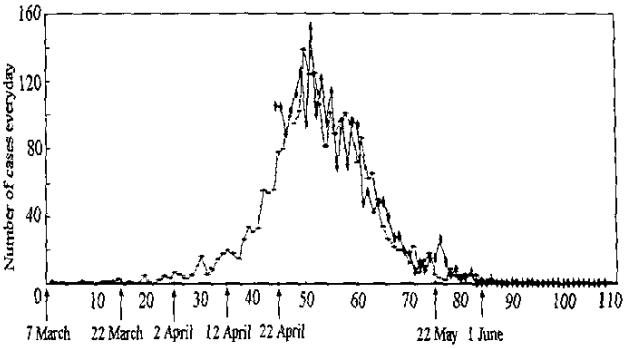

Fig.7 Comparison of open model (red) with actual data from Beijing's SARS cases (blue). Beijing's statistical data is after 22 April (see http://www.bjcdc.org/jkzl/fdpop/fddgnqk.asp, Beijing's center for disease control and prevention.)

consistent with the Beijing's SARS data before 21 April[8] to some extent, the inconsistency may come from the misreport or the actual inaccuracy. Obviously, if immediate effective and stringent control measures were taken at the beginning of April, the outbreak might have not taken place and the disaster might have been avoided. However, SARS is a new mortal infectious
disease, the basic knowledge about it is almost blank for human beings at the beginning, especially the knowledge of its biological properties and effective treatment. When thousands of doctors and nurses were endeavoring to rescue SARS patients, many of them were completely exposed to the devil — coronavirus, and were infected unconsciously at the early stage of SARS infection. By adjusting the population, we find the total number of infectious SARS patients is independent of population to some extent. The premise is the number of infected SARS cases shall be far less than the number of susceptible population.

Table I Data from simulation

	22 Mar	2 Apr	12 Apr	17 Apr	22 Apr	22 May	1 Jun
Accumulated Number	7	27	128	248	518	2454	2490
Increment everyday	0	6	19	30	77	5	0

IV COMMENTS AND CONCLUSIONS

Good consistency among mathematical formulation, simulation and actual data from Beijing indicates the correctness and rationality of the current models, which are also consistent with the related analysis in other preceding jobs[5,6,11-14]. Many interesting and important conclusions are drawn about SARS epidemics, esp. the explicit analytical mathematical formulation exhibits inspiring results, so the model can be employed in the research of SARS transmission or other similar infections.

It should be specially pointed out that the stringent control measures shall be taken in time without hesitation and delay, the earlier the stringent measures, the better of impact of controlling and eliminating infection. In the conference sponsored by WHO on 17-18 June 2003 in Malaysia, three overarching questions were addressed in the summary report

of "where do we go from here?": can SARS be eradicated? are current control measures effective? are current alert and response systems appropriately robust?[2] Obviously, severe and extensive international concerns are still focus on the effectiveness of the current control measures and alert and response systems. In the present models, system of checkpoints and health registration, isolation of SARS patients and quarantine of asymptomatic individuals turn out to be more powerful and effective.

There are undoubtedly many drawbacks existing in the present model which is still a preliminary rough model: the transition process of the value of R has been neglected, which results in a little inconsistency around the time of the institution of control measures; the expressions of many formulae are much complicated and not very convenient for further research; the probabilistic density function of incubation period and contact matrices are simplified; many social and natural factors are not taken into account, *et al.* As a matter of fact, choice of a suitable model framework is not straightforward: a variety of approaches are possible, ranging from a simple deterministic compartmental approach to a spatially explicit, individual-based simulation. A coherent estimation framework requires the use of a more sophisticated dynamical model to capture the non-linearity and spatial locality of transmission and to incorporate information on how key parameters vary between individuals.

Another important question is whether the mathematical formulation is a possible solution of epidemic differential equations. This is a very interesting and important question, which may throw light on the research on epidemic differential equations.

We are going to do more sophisticated work on more reasonable probabilistic density function (*e.g.*, incubation period, resistibility, contact degree and duration), scientific rules of adjustment of parameters, the impact of the transition process of R, other various natural and social factors that affect the spread of epidemics, and the solution of differential equations, *et al*, these will constitute our future job.

ACKNOWLEDGEMENT

Li Jinping is grateful to Li Qingshan and Liu Xingkui — doctors of attached hospital of Jinan University — for their helpful discussions.

REFERENCES

[1] http://www.who.int/csr/sars/en/. SARS: Status of the outbreak and lessons for the immediate future. *Prepared for the SARS technical briefing*, WHA 56, 20 May 2003.

[2] http://www.who.int/csr/sars/conference/june_2003/materials/report/en/. *Summary report*: where do we go from here? *WHO Global Conference on Severe Acute Respiratory Syndrome (SARS)*, 17–18 June 2003.

[3] Martin E. INFECTIOUS DISEASES: Calling All Coronavirologists. *Science*, Apr 2003; 300: 413 – 414.

[4] JSM Peiris, ST Lai, LLM Poon, *et al.* Coronavirus as a possible cause of severe acute respiratory syndrome. *The Lancet*, Volume 361, Issue 9366, 19 April 2003, pp1319-1325.

[5] Riley S, Fraser C, Donnelly, *et al.* Transmission dynamics of the etiological agent of SARS in HongKong: Impact of public health interventions. *Science*, 2003, 300, pp1961-1966.

[6] Lipsitch M, Cohen T, Cooper B, *et al.* Transmission dynamics and control of severe acute respiratory syndrome. *Science*, 2003, 300, pp1966-1970.

[7] http://www.chinacdc.net.cn/. Chinese center for disease control and prevention.

[8] http://www.bjcdc.org/jkzl/fdpop/fddgnqk.asp.Bejing's center for disease control and prevention.

[9] G. Chowell, P. W. Fenimorea, M. A. Castillo-Garsowe and C. Castillo-Chaveza, SARS outbreaks in Ontario, Hong Kong and Singapore: the role of diagnosis and isolation as a control mechanism. *Journal of Theoretical Biology*, Volume 224, Issue 1, 7 September 2003, pp1- 8.

[10] http://www.scichina.com/ZTYJ/kzzt.htm. SARS special topics. Or see *Chinese Science Bulletin*, 2003, 48, No.12-13.

[11] Shi YL. Stochastic dynamic model of SARS spreading. *Chinese Science Bulletin*, 2003 , 48(13): 1373-1377.

[12] Yang FT, Hou LJ, Han J, *et al.* Simulation analysis of Beijing's Epidemics. *Journal of System Simulation*, 2003, 15(7). (In Chinese)

[13] Chen JR, Yang FT, Zhan SY, *et al.* Processing of parameters and initial values of SARS simulation model of Beijing. *Journal of System Simulation*, 2003, 15(7). (In Chinese)

[14] Chris Dye and Nigel Gay. Modeling the SARS Epidemic. *Science*, Jun 2003; 300: 1884 – 1885.

[15] Anderson, R. M., May, R. M., Infectious Diseases of Humans: Dynamics and Control, Oxford: Oxford Univ. Press, 1991.

[16] Wang, W. D., Ma, Z. E., Global dynamics of an epidemic model with time delay, Nonlinear Analysis: Real World Applications, 2002, 3: 365-373.

[17] Daley, D. J., Gani, J. M., Epidemic Modeling: An Introduction, *Cambridge: Cambridge University Press*, 1999.

[18] Scott Gottlieb. Chinese scientists must test wild animals to find the host of SARS. *British Medical Journal*, May 2003; 326: 1109.