Jpn. J. Infect. Dis., 58, 2005

Infection Transmission Science and Models

James S. Koopman*

University of Michigan Department of Epidemiology and Center for the Study of Complex Systems, Michigan 48109, USA

SUMMARY: Infection transmission systems circulate infection through complex contact patterns related to both contact patterns and patterns of factors that affect the risk of transmission given contact. The nonlinear dynamics of infection transmission cause these patterns to make big differences in population infection levels. A science of infection transmission system analysis is needed to focus on those details that affect the control of infection transmission. This science must have a strong theoretical base because there is little chance that a dominantly data based approach not using mechanistic models of transmission will have any predictive value. The theoretical base should be built on linked transmission system models that are focused on making needed inferences for both building the theoretical base and making infection control decisions. The linking of different models is needed for a strategy of inference robustness assessment that is designed to find the model that is simple enough to effectively analyze the transmission system but not so simple that realistic violation of simplifying assumptions will change an inference. Types of models that should be used in such linked analyses include deterministic and stochastic compartmental models, discrete individual models with individual event histories but structured mass action mixing, network models that provide more detail as to who has contact with whom, and intermediate model forms such as correlation models that address some aspects of contact details while preserving the flexibility of deterministic compartmental models to structure mixing and analyze the system. While transmission system science is currently weak in regards both to its data base and its theory base, many things are now coming together that could make this science flourish. On the data side these include greater ability to detect infectious agent sequences in the environment and greater ability to sequence and genetically relate agents identified at different sites in the transmission system. On the theory sides, new model construction and model analysis methods are providing new potential to use the new sources of data. Also new parameter estimation methods provide new potential to combine models and data in effective analytic strategies.

Introduction

A science of infection transmission is needed to assess and counter the threat of newly emerging and bioterrorist spread infections as well as to reduce the high burden of disease from established infection transmission patterns. To make progress, this science must analyze infection transmission system models.

Transmission system components include infectious agents, hosts, the space in which hosts move, and environments affecting how infectious agents reproduce or survive outside of hosts. Transmission system processes include those that generate the natural history of infection and contagiousness, mechanisms by which agents leave hosts, all the modes of transmission that enable agents to survive as they transit from one host to another, the processes that lead to contact between hosts, and the evolutionary mechanisms of the agent and hosts. A movie of an actual transmission system would have a background of all the elements of the system including the dynamic movements of hosts. It would show all the infectious agents, their multiplication, and their variations in genetic structure and fitness.

Transmission system models make simplifying assumptions about the bewildering complexity of actual systems. The pattern of contacts might be simplified by just specifying an effective contact function that assumes random mixing. From that simplicity one might go on to progressively specify population groups, population movement patterns, locations where people meet, vehicles and or vectors of transmission, factors that alter the risk of transmission, and myriad other details.

Transmission system model output is a dynamic pattern of infection flows, infected and/or immune hosts, and agent diversification within population groups and sites. A flow of infection involves movement of a reproductive lineage of an infectious agent from one host population group or transmission site to another.

Transmission system model analysis involves determining how patterns of infection proliferation and flow are modified by infectious agent characteristics, host response patterns, environmental conditions and/or contaminations, patterns of contagiousness and immunity upon infection, behaviors, social structures, social events, and/or control actions. Transmission system analysis must follow the definition of the ends that the analysis is to pursue. That is because infection transmission systems have many complicated details and nonlinear dynamics that give rise to complex patterns in even simple models. These make it impossible to construct a single highly detailed model that is suitable for a wide variety of uses.

Uses for transmission system analyses

The uses of modeling include gaining general insights, generating testable hypotheses, organizing and defining what is and is not known, interpreting observations, expressing scientific theories, predicting effects of actions or events, improving communication in social decision making processes, designing studies, and analyzing data. A model analysis for one use may be robust to realistic relaxation of its simplifying assumptions while for another use it may be sensitive to those same assumptions.

A common use of modeling is to choose a course of action for infection control. When specifying such a use, the alternative actions to be explored need to be expressly defined so that the models can be constructed to incorporate the actions under consideration as well as realistic details that could

*Corresponding author: E-mail: jkoopman@umich.edu

change the choice of an action. Another use that should become increasingly important is the choice between transmission system theory alternatives. For example we might want to know the effect of environmental conditions on transmission risks or whether one particular group is sustaining transmission and disseminating it to others. For either of these two uses the inference robustness strategy to be presented later is essential. We focus mainly, however, on the question of choosing among infection control options.

Controlling infection transmission

New vaccines, chemoprophylactic agents, treatments, and disinfectants help us counter infection threats. But these tools alone are insufficient. Analyses of transmission systems are needed to indicate when and where to use these tools. Moreover it takes time to develop biologics. Consequently control tools during the early stages of emerging infections are constrained to contact tracing with quarantine, early diagnosis leading to early isolation of cases, decontamination, and contact rate reduction. These control actions are costly, difficult, and require many precise implementation decisions if they are to be effective. For example one must decide how many resources to dedicate to the speed with which quarantine is imposed, to the inclusion under quarantine of contacts that have had different classes of interaction with a case, to detection of cases with no history of contact, etc. Thus analyses of infection transmission systems are even more important during the early stage of emerging infection epidemics. To be feasible and useful early during epidemics, strong modeling traditions and methods need to be established.

Assessing the effects of actions to control infection transmission is tricky because the indirect effects of interventions are usually greater than the direct effects. That is to say, more people are protected because the chains of infection leading to them are interrupted than are protected because they received a vaccine or some other intervention. Indirect effects are especially sensitive to the nature and arrangement of contacts that transmit infection. Thus one of the key aspects of infection transmission system modeling is insuring that inferences made on the basis of model analyses are robust to realistic violation of assumptions about patterns of contact between hosts.

To prevent and control epidemics or reduce levels of endemic infection, it helps to know the population groups or environmental niches where infectious agents amplify their numbers, how they flow from one group to another through the population of interest, which sites of amplification and/or which infection flows will be affected by control actions, and how interruption of specific flows will alter the overall pattern of flows and the overall level of infection in a population. Models organize our thinking about such complex issues, provide insights about which control actions will work under different conditions, make predictions about the effects of control actions, help specify data needed to evaluate effects of control actions, and help analyze those data.

Models, data, theory, and action

Whenever we make an infection transmission control decision, we use models of transmission systems. The model may just be a mental one that provides a basis for thinking about the decision. We use the model to project what would happen under alternative actions. The more we formalize decision models and make clear their assumptions and potential fragilities, the better their predictions and our decisions will be. To explore current limitations of transmission system models and envision a path to expansion of their potential, we discuss here two dimensions that can characterize the strength of a science. These are the empirical-statistical or database dimension and the mechanistic-causal or theory dimension. The ideal science is strong in both these dimensions. Infection transmission science is currently weak in both dimensions.

The mechanistic or theory dimension captures how well processes in models relate to causal processes in the real world. A science is on the weak end when its theories do not make detailed testable predictions or when predictions made are untested by data. It is on the strong end when its detailed theories have been well tested and have made unexpected predictions that subsequent observations have confirmed.

The empirical or database dimension captures the quantity and quality of observations on the phenomena that are modeled. If model output can be repeatedly observed in the real world, such as is often the case when the phenomena are at the molecular, cellular, or individual patient levels, then the model can make accurate predictions even in the absence of mechanistic-causal theory. Sampling theory and statistical models are sufficient for good predictions given the assumption that causal processes will be the same in the future as in the past. Evidence based practice of medicine uses such theory and models. It considers the best evidence to be randomized control trials that may do little to elucidate the causal mechanisms making one therapy better than another.

Evidence based infection transmission control in populations, however, cannot proceed without strong causal mechanistic theory, however, because repeated observations at a population level are difficult and expensive. Moreover the nonlinear dynamics and complexity of infection transmission systems mean that minor differences in situations can generate big differences in outcomes. Thus, when a new epidemic appears, it is hard to specify which set of past epidemic patterns are appropriate for predicting the course of this new epidemic. Take, for example, the use of risk group focused screening programs to stop HIV transmission. In one situation there might be a core group which sustains transmission and disseminates infection to other groups. These other groups might carry on considerable transmission. This secondary group transmission, however, might eventually die out without continuing input from the core. In that case, core group screening will have broad effects on the overall population. But if the other groups have just a little more sustained transmission, they will not need continued input to sustain transmission and core group reductions will have little effect beyond the core. Nonlinear transmission systems are full of such sensitivities that require model analyses to identify. The power of non-causal models to predict differences associated with such dynamical differences will be nearly zero.

Similar subtle differences may make it difficult for observations on Marburg virus, Ebola virus, or SARS epidemics to predict the relative contributions of early diagnosis with isolation, contact tracing with quarantine, or societal actions to reduce contact in the control of an epidemic. We are not saying that we cannot learn how to control the next epidemic by studying past epidemics. We are saying that good control decisions cannot be made based on the study of the outcomes of different control actions in a series of epidemics using

models that do not capture mechanistic-causal aspects of the spread of infection.

These two dimensions are intertwined. No database can be used without theory and no theory can make good predictions without data. Inferences that appear purely statistical require the theoretical-causal assumption that the underlying causal phenomena are not changing. Moreover, data is almost always used outside the realm of statistical sampling theory in order to generalize observations on one population to another. That requires deducing that causal phenomena in the study population do not differ from those in the population of interest.

Conversely, theory requires data to test the theory. To advance a science of infection transmission systems, we need better ways to relate data and theory. One task is to develop new methods that incorporate more theory into new ways to obtain and analyze data. Another task is to construct theory in such a way that it can be tested by repeatable observations in diverse infection transmission contexts.

Until recently, most infection transmission system models used little data or theory. One reason is that mathematical analysis techniques can only be applied to simple models. Computer model analysis can proceed using more realistically complex models. But the ethos of simplicity persists and has inhibited the construction of realistically complex theory that uses data from repeatable observations on phenomena that can be generalized across transmission systems. The data epidemiologists gather are always subject to complexities that mathematical modelers have tended to ignore in order to keep things simple. Now as epidemiologists are becoming better modelers, model analyses of more realistically detailed decision processes are being pursued.

The dictum has been to keep things as simple as possible, but no simpler. When making infection control decisions, most models have probably been simpler than required to make valid decisions. But there has been a weak ethos of pursuing complexities that would show that models are simpler than possible – that their simplifying assumptions ignore realistic complexities that if incorporated into a model would change inferences about infection control decisions. We advocate here a process of inference robustness assessment to demonstrate that a model is not simpler than needed to make an inference. This approach serves both to choose between alternative control actions or between alternative theories.

The process of inference robustness assessment

Inference robustness assessment is performed by comparing an inference across different model forms. The inference may be relevant to any of the uses of models that we listed earlier. Robustness assessment is performed by determining whether an inference made by analyzing one model is changed when analyzing a different model that realistically relaxes simplifying assumptions of the first model.

To show that an inference is not robust to simplifying assumptions, four things must happen. First the inference must be explicitly stated so that it can be evaluated under different models. Second, clear criteria for meaningful differences in inferences need to be established so that it is clear when an inference is not robust. Third, the simplifying assumptions whose realistic relaxation might change the inference need to be identified. Finally, model forms that relax just the specific assumptions of interest must be identified and analyzed.

There are good reasons for saying that a model should be as simple as possible. Realistic complexity can inhibit both productive analysis and understanding. If a model has many extraneous details, then analysis to determine the effects of those details will consume resources and divert focus. As a consequence, valuable lessons that the model might have conveyed could be obscured. Moreover, if the main goal of an analysis is to provide general insights about which control actions will work under different conditions rather than to help make a specific control decision, then simplicity is called for because insights are stimulated or communicated by simple models better than by realistically complex ones.

The reasons for saying a model should be no simpler than possible are more urgent than the reasons for saying it should be as simple as possible. A model that is simpler than possible leads to bad decisions that might cost lives.

So when should one use a simple deterministic model that ignores the role of chance events captured by more complex stochastic models? When should one make the simplifying assumptions that mixing is random rather than adding realistic details about which groups mix separately from other groups or about contact networks within groups? My answer to these and similar questions is that the pursuit of simple models is always justified as long as the model analysis task is not viewed as just the analysis of a single simple model. I contend that the model analysis task should always be viewed as a robustness assessment task with analysis of a series of models incorporating various levels of realistic detail. Whenever that is the case, it is best to begin simply in the anticipation that the more thorough analysis that is possible with simple models will lead to insights that may lead to better choice of realistic details to be added to the model when concern arises that the model might be too simple to give a correct answer to a question.

One can never know for sure that a model is not too simple to give a correct answer. All models are constructed by making simplifying assumptions. To assess robustness one can relax each identified simplifying assumptions one by one starting from a baseline simple model. But it is difficult to identify all of the simplifying assumptions in simple models. Each simplifying assumption might be relaxed in multiple ways so that in fact simple models make assumptions about each of the multiple dimensions in which the simplifying assumption can be relaxed. For example, mass action contact process assumptions (1) can be relaxed by reformulating the way that the number of individuals in a population affects the number of contacts each individual makes, by subdividing compartments across multiple dimensions, or by formulating a multitude of different processes that could link one individual to another in a network. It might seem that the mass action model makes a simple assumption. But in this light it makes a whole series of different unrealistic assumptions. Moreover, even if one is able to identify all the assumptions that allow for different possible ways of relaxing the mass action assumptions, checking out each assumption by relaxing it alone does not provide assurance that none of the relaxed assumptions are problematic. One cannot be sure that relaxing any set of two simplifying assumptions will not change an inference when each relaxation alone did not.

How can effective decisions be reached when one can never be sure that one has not failed to identify a set of assumptions that have led to a bad decision? The answer is in the social process of science and the social process of infection control. On the science side, critical colleagues might see holes in an inference missed by an original model creator. Likewise, vested

parties in control decisions are likely to see issues to address that the scientists originally constructing a model have missed. To serve both these ends, modeling strategies should be used that promote a scientific discourse where epidemiologist, modelers, and others find it easy to modify and reanalyze the models that their colleagues have constructed and analyzed.

Model forms to be used in inference robustness assessment

Multiple model forms can represent infection transmission systems. We discuss here only a few of these. Sometimes a whole series of different model assumptions can be identified across two different model forms. Comparisons with fewer assumption differences generate more knowledge about the effects of assumptions and thus contribute more both to the advance of theory and to effective decision making. Further discussion of this can be found elsewhere (2,3). The model dimensions we discuss here are not exhaustive. They include continuous populations vs. discrete individuals, deterministic vs. stochastic, compartmental versus unique individual, mass action contact versus population pattern contact versus individual network contact.

Some traditions use differential equations to model population groups characterized by their state of infection and immunity and by how they interact with other population groups. These are deterministic compartmental models that model continuous populations rather than discrete individuals. They assume that the population in each compartment can be broken into an infinite number of individuals. They are the basis of a couple classical infection transmission modeling texts (4,5). Deterministic means that the model generates a single pattern of compartment sizes with no variation being generated by chance. Compartmental means that homogeneous population groups are modeled where all population in a compartment is homogeneously identical.

A variety of approaches to formulating contact patterns are possible within this tradition. Different contact formulations within this tradition generate different total numbers of contacts as the population sizes in mixing groups change (6,7). Metapopulation models have individuals migrating between groups to generate patterns of contact (8). Structured mixing models use a statistical mechanics approach to location of population. They never specify where population is, only the chances that it is in one place or another (9). Because deterministic compartmental models are of continuous populations, they have a harder time dealing with contact structure issues that arise at the individual level such as bias in the probabilities that two contacts of one person are in contact with each other. Contact pattern issues of this sort can have large effects on transmission dynamics. A correlation model approach has been devised to address this sort of issue within the context of differential equation models (10). However, many situations where networks of contacts affect transmission dynamics are difficult to address when modeling continuous populations rather than discrete individuals.

Stochastic compartmental models assume that population sizes can only take on discrete integer values. They can formulate the same contact patterns that deterministic compartmental models can. As one relaxes the infinite population size assumption by transiting from a continuous population model to a discrete integer population model, one is usually forced to relax the deterministic behavior assumption and model a stochastic process.

Discrete individual models do not make homogeneous compartmental assumptions. They allow each individual to experience unique effects of causal model parameters so that what was a single valued parameter in compartmental models becomes a distribution of parameter values in discrete individual models.

The number of model assumptions that can be relaxed by transiting from one model form to another is almost infinite. A few illustrative assumptions and their relevant transitions are listed in Table 1.

Data analysis for inference robustness assessment

One way that transmission system science is gaining strength is through the development of new ways to use transmission system models in the analysis of data. Transmission system models assume nonlinear dynamics where population effects are not just the sum of individual effects. Data analysis using such models presents some difficult challenges. Least squares parameter fitting approaches require such extensive searches of parameter space that they become computationally challenging. When one uses approaches like

Table 1. Illustrative transitions between model types for the purpose of relaxing simplifying model assumptions

Simpler model type	More complex model type	Assumptions relaxed
Deterministic compartmental model with random mixing	Stochastic compartmental model with random mixing	Infinite population size in each compartment
Stochastic compartmental model with random mixing	Stochastic compartmental model with group structured mixing	Equal probability of contacting different classes of individuals
Stochastic compartmental model with group structured mixing	Individual event history model with group structured mixing and a distribution of individual transmission risks	Fixed effects for transmission risks become random effects
Individual event history model with mass action assumptions	Network model with same macro contact patterns and random linkage process	Instantaneous and thorough mixing of population after each contact
Network model with same macro contact patterns and random linkage process	Network model with same macro contact patterns and biased linkage process	No selectivity for partners to whom long term linkages will be made
Network model with same macro contact patterns and biased linkage process	Agent based model where characteristics of contacts are screened according to past history with such individuals	Markov assumption that past experiences do not affect linkage parameter values

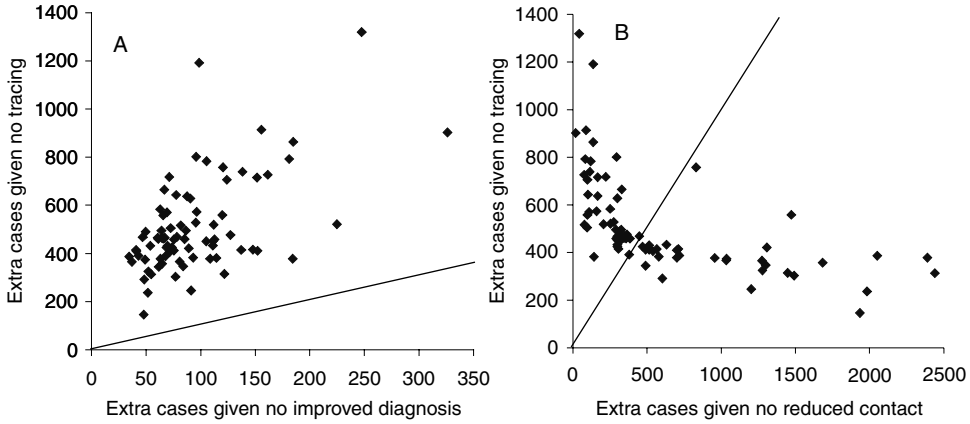

Fig. 1. Robustness assessment of an analysis that fit data from the Beijing SARS epidemic to determine the relative effects of contact tracing with quarantine, early diagnosis and isolation of cases without a history of contact, and contact reductions in controlling the epidemic.

bootstrap methods to get a handle on the variance in such estimates generated by sampling issues, the computational challenges are increased. Markov Chain Monte Carlo (MCMC) methods of analysis have produced promising advances (11). They, however, can be even more computationally challenging. The challenges for both methods are compounded by the need for robustness assessment. Any single set of parameter estimates from a least squares approach or any set of distributions of parameter values from an MCMC approach must make the assumptions of the underlying transmission system model. The parameter estimates made might not be robust to realistic violation of assumptions in the transmission model. Thus, just as the pure model analysis discussed in previous sections should use robustness assessment methods, data analysis should also use such methods.

One approach is to construct a model that captures realistic complexities of a transmission system by setting some parameter values while allowing the parameters to be estimated to vary. An example using a least squares approach to fitting parameters of a deterministic compartmental model is presented for SARS in Beijing in the accompanying figure. A series of realistic model dimensions were fixed across a range that allowed for fitting other parameters such that the average difference between model generated and observed data points was less than 20%. Then using the parameters estimated for each fit, the implications of parameter estimates were assessed for inferences about the relative benefit that was gained during the epidemic from contact tracing with quarantine of contacts, early diagnosis and isolation of cases without previously identified contact histories with infected individuals, and contact rate reduction. Each point in the figure represents a set of parameter values that includes the full range of each parameter value that still allows for fitting the observations. As can be appreciated in panel A of Figure 1, the inference that contact tracing with quarantine generated more benefit than early diagnosis with isolation is robust in that this inference holds across a full range parameter sets. Inferences about the relative benefits of contact rate reductions and the other interventions were not robust, as seen from the data points in panel B. A combination of identifiability issues and transmission system sensitivities account for this lack of robustness.

A vision of the future of transmission system science

I believe that transmission system science is entering an era where it will move up both database and theory strength scales to become a powerfully predictive science. The key elements in this advance will be new sources of data, more involvement of field and theoretical epidemiologists in transmission system modeling as software makes modeling easier for them, and continuing advances in the use of transmission system models for data analysis.

One key data advance is the increasing ability to sequence genomes from infectious agents isolated in different parts of the transmission system at different times. Because transmission generally fixes variation in sequences that arise during infectious agent proliferation within a host, transmission system models generate expected patterns of genetic relationships between organisms isolated from different parts of the system. The observed genetic relationships can thus be used to estimate transmission system model parameters.

Another key data advance is the identification of infectious agents in environmental samples. Such identifications in representative parts of the transmission system can be pursued more cheaply than agent identifications from patients in representative parts of the transmission system. Moreover, new technology for this task, such as the laboratory on a chip technology presented in another presentation in the meeting (12), will lower the costs and make more thorough sampling for epidemiological analysis possible.

As software, environmental agent identification, and agent sequence analyses improve, more specific hypotheses about the environmental conditions and personal behaviors that affect transmission via various modes will be investigated. This, together with experiences from robustness assessments using the new software, will lead to stronger theory on which to base model construction. The stronger theory will in turn allow for more extraction of information from the data in terms of more focused and informative assessments of inference robustness.

To realize this vision, computer scientists, mathematicians, biochemists, industry, and above all epidemiologists must work together. I believe that an inference robustness assessment focus is a good way to insure productive collaboration in these efforts.

REFERENCES

1. Koopman, J. S. (2005): Mass action and system analysis of infection transmission. *In* Cuddington, K. and Beisner, B. E. (eds.), Ecological Pardigms Lost: Routes to Theory Changes. Academic Press.
2. Koopman, J. S. , Jacquez, G. and Chic, S. E. (2001): New data and tools for integrating discrete and continuous population modeling strategies. Ann. New York Acad. Sci. 954, 268-294.
3. Koopman, J. S. (2004): Modeling infection transmission. Ann. Rev. Public Health, 25, 303-326.
4. Anderson, R. M. and May, R. M. (1991): Infectious Diseases of Humans: Dynamics and Control. Oxford Universtiy Press, Oxford.
5. Diekmann, O. and Heesterbeek, J. A. P. (2000): Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation. Wiley, Chichester.
6. McCallum, H., Barlow, N. and Hone, J. (2001): How should pathogen transmission be modelled? Trends Ecol. Evol., 16, 295-300.
7. Riggs, T. W. and Koopman, J. S. (2004): A stochastic model of vaccine trials for endemic infections using group randomization. Epidemiol. Infect., 132, 927-938.
8. Grenfell, B. and Harwood, J. (1997): (Meta)population dynamics of infectious diseases. Tree, 12, 395-399.
9. Jacquez, J. A., Simon, C. P. and Koopman, J. S. (1989): Structured mixing: heterogeneous mixing by the definition of activity group. p. 316-349. *In* Castillo-Chavez, C. (ed.), Mathematical and Statistical Approaches to AIDS Epidemiology. vol. 83. Springer-Verlag, Heidelberg.
10. Keeling, M. J., Rand, D. A. and Morris, A. J. (1997): Correlation models for childhood epidemics. Proceedings of the Royal Society of London - Series B: Biological Sciences. 264 (1385), 1149-1156.
11. O'Neill, P. D. (2002): A tutorial introduction to Bayesian inference for stochastic epidemic models using Markov chain Monte Carlo methods. Math. Biosci., Nov-Dec 180, 103-114.
12. Suzuki, K., Yamamoto, K. and Yoshikura, H. (2005): International symposium on infectious agent transmission model building - focusing on assessment of risk to communities. Jpn. J. Infect. Dis., 58, S1-S2.