EDITORIAL

Complexity, simplicity, and epidemiology

Neil Pearce1,2* and Franco Merletti2

Accepted 4 January 2006

Keywords Epidemiology, complexity, risk factors, social theory

Introduction

It is difficult, nowadays, to open a popular science magazine, or a leading science journal, without reading about complexity, the approach to science that is expected to ‘define the scientific agenda for the 21st century’.1 Complexity theory is influencing fields as diverse as physics,2 cosmology,3 chemistry,4 geography,5 climate research,6 zoology,7 biology,8 evolutionary biology,9 cell biology,10 neuroscience,11 clinical medicine,12 management,13 and economics.14 However, it has to date had relatively little influence on the theory and practice of epidemiology.15 In this paper we review the basic concepts of complexity theory and discuss their relevance to epidemiology.

Complexity

It should be stressed that although many phenomena are complex,15 the concept of ‘complexity’ is more specific. Complexity is the study of complex adaptive systems. These have been defined as ‘a collection of individual agents with freedom to act in ways that are not always totally predictable, and whose actions are interconnected so that one agent’s actions changes the context for other agents’.

Such systems include living cells, the brain, the immune system, the financial markets, ecosystems, and human populations. They are complex in the sense that there are a great many apparently independent agents interacting with each other, but the richness of these interactions allows the system as a whole to undergo self-organization.1 They are also characterized as involving non-linearity and feedback loops in which small changes can have striking effects that cannot be understood simply by analysing the individual components.17 The whole is more than the sum of its (reductionist) parts. Such complex systems can exist on a number of different levels from the subatomic through to the individual level, the population level, and beyond.18

The most striking example of a complex self-organizing system is life itself, not only in terms of individual organisms but also in evolutionary terms—organisms adapt to each other through evolution into a finely tuned ecosystem. Similarly, various populations have evolved traditional ways of life that are now responding to the changes brought by the industrial revolution, colonization, and globalization.19

Thus, a key feature of such complex systems is that they are adaptive. They do not just passively respond to events, but they reorganize themselves into a new equilibrium in response to events.1 The brain reorganizes itself to learn from experience, species evolve to achieve a new ecosystem in response to events such as climate change or meteor strikes, and populations evolve in response to economic and social changes often while retaining their ‘traditional’ cultures in a new form and context.

Such a dynamic equilibrium is not always achieved—species become extinct, populations and cultures are extinguished, and financial markets go into freefall—but new emergent forms of self-organization arise from ‘the edge of chaos’1 to take their place. The new forms of organization that may arise are often unpredictable because small changes in the initial conditions may produce large changes in the final equilibrium state that is achieved. However, although the details may be unpredictable, the general shape of the new forms of organization may be relatively predictable and simple. For example, small changes in the initial conditions may have drastically changed the evolutionary story, but the superficial ‘forms’ of evolution are likely to have been similar despite the different routes involved—it is likely that something resembling birds (with wings, feathers, etc.) would have evolved to fill an ecological niche even if the evolutionary pathway had been markedly different.20

This illustrates another key feature of the complexity theory that what appears chaotic and unpredictable at one (usually lower) level may be relatively simple and stable at another (usually higher) level.18 No one would attempt to predict the weather from measurements of individual molecules—all you would see is noise—but such weather systems can be extremely simple and predictable when observed at the appropriate level,21 and the concept of ‘climate’ is a summary of the broad patterns of weather that may be more predictable, although even the climate system may be sensitive to small perturbations and may change over time. Thus ‘nature can produce complex...
structures even in simple conditions, and can obey simple laws even in complex situations. Climate may be very complex and difficult to predict on a day-to-day basis, but winter regularly follows summer and El Nino occurs at semi-predictable intervals.

These five concepts—self-organization, adaptation, upheavals at the edge of chaos, the unpredictability of the effects of small changes in the initial conditions, and the existence of simplicity at some levels while ‘chaos’ exists at others—form the fundamental concepts of complexity. In fact, such ideas are not new but are in part based on dialectical methods of thinking that have a long history in science but have been rediscovered and adapted in the past few decades.

These concepts did not fit conventional thinking in physics, biology, or economics when they first (re)appeared. They ask different questions. For example, until recently modern economic theory was (and largely still is) confined to the theoretical study of free markets that are in perfect equilibrium. It has at best a tenuous connection with the messiness of the real world in which historical conditions, political decisions, monopolies, etc. set the boundaries within which markets operate. Once the existence of such phenomena is recognized, many (non-linear) scientific questions arise, which simply do not get asked, and cannot be tested, under standard (linear) economic theory.

Epidemiology

So what do such concepts have to do with epidemiology? The most obvious connection is that the health of a population can be viewed as a complex adaptive system. A population is not just a collection of individuals; rather, each population has its own history, culture, and socioeconomic structures, which survive despite massive global economic change while at the same time being affected and shaped by such change. The health of a population is shaped by, and shapes, the sociocultural context in which the population lives. Thus, although the occurrence of disease can be studied at many different levels, including ecosystems, populations, individuals, and molecules, it has been argued that the population level is fundamental for epidemiology. There are clearly exceptions to this; e.g. the ecosystem level is crucial when considering the long-term health consequences of climate change and the individual and molecular levels of analysis, and the interactions between the various possible levels of analysis are also important.

However, the population level is generally fundamental in public health terms, since it defines the public health problems that should be addressed. Furthermore, it is also often fundamental in scientific terms since some scientific problems can be best understood at the population level and cannot be reduced to the individual or molecular levels.

As noted above, there are very few examples of the use of the complexity theory in epidemiology, but there are many examples of epidemiological problems for which the complexity theory is relevant. In particular, although a focus on the population level, and the sociocultural context, does not necessitate the use of the complexity theory, it makes its value and potential more apparent. Therefore, in this section, we discuss examples of the relevance of the complexity theory to epidemiology.

Communicable disease

To date, the complexity theory has received the most application in epidemiology with regard to research into communicable disease. The interactions between the variables that determine the transmission of infections in populations are often complex and non-linear. Network theory can capture the diversity of human contacts that underlie the spread of diseases such as SARS and can lead to different predictions, and different interventions, than those generated by more orthodox ‘compartmental’ models in which each person in a population has an equal chance of spreading the disease to everyone else. In particular, orthodox theory predicts that all such outbreaks should spark large-scale epidemics, but this is often not the case.

Koopman argues that appropriately modelling the transmission of infections requires the use of computer models that vary from deterministic models of continuous populations to models of dynamically evolving contact networks between individuals. He argues that ‘much more is needed to understand the determinants of infection flows through a population in the manner that science has helped understand the determinants of weather and ocean currents flows’. Such complex models can answer questions such as ‘which populations or places deserve concentrated intensive surveillance or control efforts like quarantine, chemoprophylaxis, symptomatic treatment, vaccination or decontamination?’ or ‘Should control be sought with interventions directed to the entire population or will tracing and quarantine be more productive?’

Similarly, Auld considers a dynamic model of risky behaviour in the midst of an epidemic and shows that the effect of policy interventions, such as preventative vaccines, may depend on whether the intervention was anticipated.

Thus a complexity-based approach to communicable disease involves quite different types of scientific question than asking ‘does virus A cause disease B?’ or ‘what risk factors are associated with the transmission of infection?’ The latter questions can be answered using straightforward methods (e.g. the relative risk of transmission of infection in those exposed compared with those not exposed to a particular factor), and will produce findings that are in principle generalizable, but may in practice be insufficient for the control of infection in a particular population.

For example, studies of the effects of climate change on the spread of malaria may involve models based on factors such as the human-biting rate of mosquitoes, human susceptibility, mosquito susceptibility, daily survival probability of the mosquito, and the incubation period of the parasite; these underlie the spread of diseases such as SARS and can lead to different predictions, and different interventions, than those generated by more orthodox ‘compartmental’ models in which each person in a population has an equal chance of spreading the disease to everyone else. In particular, orthodox theory predicts that all such outbreaks should spark large-scale epidemics, but this is often not the case.

Koopman argues that appropriately modelling the transmission of infections requires the use of computer models that vary from deterministic models of continuous populations to models of dynamically evolving contact networks between individuals. He argues that ‘much more is needed to understand the determinants of infection flows through a population in the manner that science has helped understand the determinants of weather and ocean current flows’. Such complex models can answer questions such as ‘which populations or places deserve concentrated intensive surveillance or control efforts like quarantine, chemoprophylaxis, symptomatic treatment, vaccination or decontamination?’ or ‘Should control be sought with interventions directed to the entire population or will tracing and quarantine be more productive?’

Similarly, Auld considers a dynamic model of risky behaviour in the midst of an epidemic and shows that the effect of policy interventions, such as preventative vaccines, may depend on whether the intervention was anticipated.

Thus a complexity-based approach to communicable disease involves quite different types of scientific question than asking ‘does virus A cause disease B?’ or ‘what risk factors are associated with the transmission of infection?’ The latter questions can be answered using straightforward methods (e.g. the relative risk of transmission of infection in those exposed compared with those not exposed to a particular factor), and will produce findings that are in principle generalizable, but may in practice be insufficient for the control of infection in a particular population.

For example, studies of the effects of climate change on the spread of malaria may involve models based on factors such as the human-biting rate of mosquitoes, human susceptibility, mosquito susceptibility, daily survival probability of the mosquito, and the incubation period of the parasite; these underlie the spread of diseases such as SARS and can lead to different predictions, and different interventions, than those generated by more orthodox ‘compartmental’ models in which each person in a population has an equal chance of spreading the disease to everyone else. In particular, orthodox theory predicts that all such outbreaks should spark large-scale epidemics, but this is often not the case.

Koopman argues that appropriately modelling the transmission of infections requires the use of computer models that vary from deterministic models of continuous populations to models of dynamically evolving contact networks between individuals. He argues that ‘much more is needed to understand the determinants of infection flows through a population in the manner that science has helped understand the determinants of weather and ocean current flows’. Such complex models can answer questions such as ‘which populations or places deserve concentrated intensive surveillance or control efforts like quarantine, chemoprophylaxis, symptomatic treatment, vaccination or decontamination?’ or ‘Should control be sought with interventions directed to the entire population or will tracing and quarantine be more productive?’

Similarly, Auld considers a dynamic model of risky behaviour in the midst of an epidemic and shows that the effect of policy interventions, such as preventative vaccines, may depend on whether the intervention was anticipated.

Thus a complexity-based approach to communicable disease involves quite different types of scientific question than asking ‘does virus A cause disease B?’ or ‘what risk factors are associated with the transmission of infection?’ The latter questions can be answered using straightforward methods (e.g. the relative risk of transmission of infection in those exposed compared with those not exposed to a particular factor), and will produce findings that are in principle generalizable, but may in practice be insufficient for the control of infection in a particular population.

For example, studies of the effects of climate change on the spread of malaria may involve models based on factors such as the human-biting rate of mosquitoes, human susceptibility, mosquito susceptibility, daily survival probability of the mosquito, and the incubation period of the parasite; these underlie the spread of diseases such as SARS and can lead to different predictions, and different interventions, than those generated by more orthodox ‘compartmental’ models in which each person in a population has an equal chance of spreading the disease to everyone else. In particular, orthodox theory predicts that all such outbreaks should spark large-scale epidemics, but this is often not the case.

Koopman argues that appropriately modelling the transmission of infections requires the use of computer models that vary from deterministic models of continuous populations to models of dynamically evolving contact networks between individuals. He argues that ‘much more is needed to understand the determinants of infection flows through a population in the manner that science has helped understand the determinants of weather and ocean current flows’. Such complex models can answer questions such as ‘which populations or places deserve concentrated intensive surveillance or control efforts like quarantine, chemoprophylaxis, symptomatic treatment, vaccination or decontamination?’ or ‘Should control be sought with interventions directed to the entire population or will tracing and quarantine be more productive?’

Similarly, Auld considers a dynamic model of risky behaviour in the midst of an epidemic and shows that the effect of policy interventions, such as preventative vaccines, may depend on whether the intervention was anticipated.

Thus a complexity-based approach to communicable disease involves quite different types of scientific question than asking ‘does virus A cause disease B?’ or ‘what risk factors are associated with the transmission of infection?’ The latter questions can be answered using straightforward methods (e.g. the relative risk of transmission of infection in those exposed compared with those not exposed to a particular factor), and will produce findings that are in principle generalizable, but may in practice be insufficient for the control of infection in a particular population.
occurred primarily because of the arrival of infectious diseases to which the Maori had no natural immunity. However, a more careful analysis of the history of colonization throughout the Pacific reveals that the indigenous people mainly suffered major mortality from imported infectious diseases when their land was taken, thus disrupting their economic base, food supply, and social networks. The population context was as important as the exposure itself.

Similarly, McKean has documented the dramatic decline in mortality during the past century from the ‘diseases of poverty’ that were dominant in the 19th century—particularly infectious diseases, respiratory diseases, and accidents, and has argued that the decline can be attributed mainly to improvements in nutrition. Alternatively, it has been argued that specific public health interventions on factors such as housing and urban congestion actually played the major role.

Debate continues regarding the explanations for the decline in mortality, but whatever the explanation, it is clear that the socioeconomic context played a major role.

Consideration of the specific population context for infectious disease yields knowledge that is highly specific but also relevant to other populations and other contexts. Thus, the experience of New Zealand (Aotearoa) and the UK in the 19th century, makes it less surprising that in the late 20th century the countries of Eastern Europe experienced the largest sudden drop in life expectancy that has been observed in peacetime in recorded human history with a major rise in ‘forgotten’ diseases such as tuberculosis and cholera as well as in cardiovascular and other alcohol-related diseases.

Non-communicable disease

Complexity theory has been used to study the occurrence of non-communicable disease at the clinical (individual) level. Many healthy states represent complex equilibria, whereas disease states represent a breakdown of self-organization and a collapse into less complex dynamics. Illness arises from dynamic interaction within and between self-adjusting systems not from a failure of a single component.

However, there have been relatively few attempts to explicitly use complexity theory to study non-communicable disease occurrence at the population level, although the history of public health is full of examples to which complexity theory is relevant.

In fact, just as social and economic conditions can explain why some people, and not others, are exposed to infections, they also are relevant to exposures to risk factors for non-communicable disease. Why cannot transmission of tobacco smoking in the population be modelled using similar techniques to those for modelling the transmission of infection? Both are affected strongly by socioeconomic circumstances, by exposures within households and families, by social networks, and by the intensity of the exposure. Both result in exposure distribution patterns that are non-random and involve complex adaptive systems. Any meaningful public health intervention on tobacco must also consider why manual workers smoke more than non-manual workers and find it more difficult to give up, why smoking is increasing among women in many countries, and why most physicians have responded to the epidemiological evidence and given up smoking whereas nurses continue to smoke in great numbers. More generally, people live within networks that have a profound effect on their health ‘choices’. Such ‘lifestyle choices’ may be in a relatively stable equilibrium and cannot be changed simply by changing one component of a complex system, e.g. by giving health education advice while ignoring the social circumstances of those ‘receiving’ the advice. Similarly, it is no accident that environmental hazards are not randomly distributed with respect to ethnicity and social class.

Some risk factors operate relatively proximate (downstream) to the final event while others operate at a greater ‘distance’ (upstream). The more proximate the exposure is to the event, the greater is the linear impact; ‘upstream’ exposures may have just as great an impact, but the effects may be non-linear and less predictable. Thus, such research does not always lead to high predictability, but the lower emphasis on prediction carries with it a greater emphasis on understanding of the processes being observed, rather than simply having a ‘black box’.

Conclusions

So what are the implications of complexity theory for epidemiology?

Complexity theory emphasizes the shortcomings of naïve reductionism. People are not just random collections of cells or molecules, and populations are not just random collections of individuals. Complex adaptive systems have a ‘life’ that is more than the sum of their component parts. Understanding brain function requires not only a knowledge of its constituents but also an understanding of the systematic context in which they operate. Risk factors for disease do not operate in isolation but occur in a particular population context. Individual ‘lifestyle’ can only be understood in the historical, cultural, and social context in which it occurs.

Complexity theory also emphasizes the importance of the concept of ‘levels of analysis’. What is chaotic at one level may be simple at another level, but to obtain useful knowledge one must focus on the appropriate level. We do not need to understand what is going on at the molecular level in order to send a rocket to the moon, nor do we need to focus on the molecular level to achieve improvements in public health; in fact, in both instances a sole focus on the molecular level would make such an enterprise impossible.

Complexity theory also emphasizes the need to develop new methods that are appropriate for the problem under study. How can you test theories about a complex adaptive system using ‘standard’ epidemiological methods? Usually you can not. Complexity theory does not fit with standard approaches to epidemiology any more than it fitted with the standard approaches to other sciences until recently. Much of modern epidemiological thinking has involved studying the effects of exposures in individuals. Complexity theory emphasizes that the populations epidemiologists study are not just collections of individuals and that the population context is not just noise but may in some instances be fundamental. Once this is recognized, a whole new set of scientific questions arise that span both epidemiology and demography and involve quite different methods from the usual epidemiological techniques.
Thus, if we are not to be ‘prisoners of the proximate’ then it will be necessary to develop new epidemiological methods that are more appropriate for the complexity of population health.53 When we are studying ‘downstream’ ‘proximate’ factors our standard methods will continue (in general) to work well, but as attention moves ‘upstream’ to the population level, modern epidemiological methods will become increasingly inappropriate, and new methods will need to be developed.55 In some instances this will involve developments of existing methods that take into account complexity and multiple levels of analysis (e.g. multilevel methods, Bayesian approaches, causal graphs, etc.), whereas in other instances it will involve the development of completely new methods or the adaptation of methods from other disciplines. There is nothing particularly unusual in this: all sciences develop new methods in response to new problems. As McMichael58 notes ‘who had heard of a case–control study or a multivariate personalised risk score this time last century?’. The appropriateness of any research methodology depends on the phenomenon under study: its magnitude, the setting, the current state of theory and knowledge, the availability of valid measurement tools, and the proposed uses of the information to be gathered, as well as the community resources and skills available and the prevailing norms and values at the national, regional, or local level.55

Complexity research involves non-linearity and ‘feedback loops’, which cannot be neatly summarized in a 2 × 2 table. Thus, the new methods that will need to be developed will look less like a randomized controlled trial—you can not do a cohort study of climate change unless you have two planets—and more like complex observational research such as evolutionary biology or cosmology. It will involve greater use of methods such as causal graphs,59,60 and other methods than can be used to model complex adaptive systems.

A complexity-based approach produces findings that are more specific to the population under study, but which have more direct public health relevance and validity. Paradoxically, the model that is utilized in a complexity approach may, therefore, be more generalizable to other populations. ‘Local’ research that is grounded in a particular population is more likely to produce findings that address universal themes and issues than is research that attempts to strip away the population context.61

As with any new theoretical approach, complexity theory is not a panacea and has the potential for misuse.15 Nevertheless, it also has considerable potential to assist epidemiology to address the major global public health problems of the 21st century. There is an old saying in the army that ‘the generals are always ready to fight the last war’. In other words, generals usually use methods and strategies that were appropriate in the last major war but may be completely inappropriate in a new context. Are we going to continue to use the epidemiological methods of the 20th century to address the scientific and public health problems of the 21st century? If we wish to bring epidemiology into the 21st century, complexity theory is likely to play an important role.

Acknowledgements

N.P.’s work on this paper was supported by the Progetto Lagrange, Fondazione CRT/ISI and the Health Research Council of New Zealand. We thank Dr Rodolfo Saracci for his comments on the draft manuscript.

KEY MESSAGES

- Complexity theory is influencing many diverse fields of science but has had little influence to date on the theory and practice of epidemiology.
- Complexity is the study of complex adaptive systems.
- The health of a population can be viewed as a complex adaptive system.
- The key concepts of complexity theory are self-organization, adaptation, upheavals at the edge of chaos, the unpredictability of the effects of small changes in initial conditions, and the existence of simplicity at some levels while chaos exists at others.
- To date complexity theory has received the most application in epidemiology with regard to communicable disease, but there is considerable potential for its application to the study of non-communicable disease.
- It will be necessary to develop new epidemiological methods that are more appropriate for addressing the complexity of population health.

References