Chaos, Criticality, and Public Health

Robert E. Fullilove, EdD, Jennifer Choe Edgoose, MD, and Mindy Thompson Fullilove, MD
New York, New York

Historically, models describing health and disease processes have been grounded in causation, and, more precisely, in an implicit faith that causal models make accurate predictions about the causes and effects of disease. The complexities of modern life, however, have led us to rethink "reductionist" approaches to epidemics, health, and health care. Recent attempts to understand the dynamics of large, interactive systems, particularly those that demonstrate what has been described as "self-organized criticality," offer potentially more robust descriptions of these phenomena. One key feature of these models is that large-scale perturbations may be generated by events that are, comparatively, quite small. Application of these models to urban public health settings suggests that the collapse of inner-city communities, such as the Bronx, New York, and accompanying epidemics of crime, drug abuse, and human immunodeficiency virus (HIV) disease, may be traced to much smaller events (eg, decisions to reduce fire services). These models also suggest, however, that effective interventions may be made that are, themselves, comparatively modest.

BURDEN OF DISEASE

> Dealing with the social causes of disease and responding to its medical effects are not mutually exclusive. We should do both. People are already burdened by poverty and lack of education should not also carry a disproportionate share of illness. Americans live in a society that tolerates great disparities in wealth and privilege. The social and political costs are enormous. We are now learning that the medical costs are also very high.[1]

The health of the public might be described, crudely, as the relationship between the level of illness in a given community and the availability of health care. If the demand for such care exceeds the supply, it is reasonable to assume that the overall health of the community will decline. What happens to the health of the public, however, in a setting in which the demand for health care vastly exceeds the supply?

Shoemaker and colleagues[2] describe such a setting. At the King/Drew Medical Center in Los Angeles, a public hospital that serves a poor, predominantly minority population, 82% of hospital admissions are through the emergency department. The authors describe this pathway to hospital care as "a bedeviling misuse of the emergency system as originally conceived." Dire consequences have ensued. In 1991, the emergency room at King/Drew was closed to incoming patients 67.5% of the time "because of saturation of the facilities." Due to shortages of hospital beds and treatment facilities, 60% of the 28,150 gunshot victims treated at King/Drew over a 14-year period were treated as outpatients. The authors further observe that, ultimately, such conditions increase morbidity and mortality among hospital patients simply because treatable ailments cannot be managed efficiently or effectively.

With a hospital-based health-care system in "overload" such as this, one might well speculate about the ripple effects such conditions will have in the community at large. If the disease burden borne by community members increases, a variety of "disaster scenarios" can be envisioned in which, say, large numbers of people die in a short space of time because an otherwise manageable medical emergency gets out of hand.

Is this, however, a reasonable, valid extrapolation from the current status of health-care delivery in South Central Los Angeles or in any similar urban setting? Do current models of public health

---

From the School of Public Health and the College of Physicians and Surgeons, Columbia University, and the New York State Psychiatric Institute, New York, New York. Requests for reprints should be addressed to Dr Robert E. Fullilove, Associate Dean, Columbia University School of Public Health, 600 W 168th St, New York, NY 10032.

and disease control suggest that such a collapse is either possible or eminent?

THEORIES OF CAUSALITY IN PUBLIC HEALTH

Historically, models describing health and disease have been grounded in causation and, more precisely, in an implicit faith that casual models "work." In the 19th century, germ theory not only established that pathogens "caused" disease, it suggested that maintaining the health of the public involved the performance of two basic tasks: finding pathogenic causes of disease and applying appropriate medications. In 1854, Dr Snow, the founder of modern epidemiology, refined this model when he deduced that an outbreak of cholera in London could be traced to the use of the local water pump.[3] By further inferring that cholera was transmitted via the water supply, Snow helped advance the notion that the science and practice of public health should be dedicated to the identification of the causes and effects of disease.

The complexities of modern life, however, have led us to rethink this "reductionist" approach to epidemics, health, and health care. We have not abandoned our beliefs that effects follow from causes, but the models we use to understand the world are increasingly focusing on the myriad ways in which complex systems interact. With vast technological advances in communication and transportation, the members of the human family have become interconnected in ways that were previously unthinkable. Voice contact with virtually any individual on the planet is no more than a millisecond away, and no location is more than a few hours travel from any other. As the strands of the human ecological web continue to be woven into an even finer mesh, the patterns of human interactions become ever more complex. These interactions have consequences that affect the physical, intellectual, and emotional health of the individual. Linear models of causation provide only a partial description of this complexity, and other models may prove to be more useful. In the sections that follow, we present the elements of such a model that is rooted in our efforts to understand complexity, chaos, and the functioning of large interactive systems.

THE DYNAMICS OF CRITICALITY

While linear dynamics provides a useful model for predicting the outcomes of interactions among elements in simple, closed systems, it is unfortunately the case that in complex interactive systems, the problems of prediction can become intractable.[4] One of the earliest examples of our inability to subject natural phenomena to deterministic models was provided by meteorologist Edward Lorenz. Lorenz demonstrated that weather is influenced by so many small phenomena (arguably every dynamic action on the planet) that any prediction about the future state of the weather deteriorates rapidly in accuracy the further we get from the initial conditions on which our predictions are made. This sensitivity to initial conditions is so great that the fluttering of a butterfly in Rio may be said to affect the weather in Texas.[5] In 1975, James York coined the term "chaos theory" to describe what has now become a family of theories that concern the fluctuations between stable and unstable states in dynamic systems.[6] Although chaos may appear to be merely a chic "new science," its attraction lies in its practical application to a wide range of phenomena from earthquakes[7] to urban evolution[8] to epidemics.[9]

This evolution from stable to unstable states is increasingly being recognized as a characteristic of the structure and organization of many complex interactive systems. Bifurcation theory, for example, is among the important models for describing nonequilibrium systems. Prigogine and Stengers[10] describe how unstable systems, fluctuating widely around a given state of equilibrium, can suddenly jump to a new, more stable state. They described this process as "self-organizing."

Bak and Chen[11] suggest the existence of some interactive systems in which the oscillation between stable and unstable states includes periodic "catastrophic" events, such as stock market crashes or epidemic outbreaks. One of the defining elements of such systems is that they "naturally evolve to a critical state in which a minor event starts a chain reaction that can affect any number of elements in the system." These systems are said to have "self-organized criticality."

The terms "critical state" and "chain reaction" are deliberately chosen to evoke images of nuclear fission in which a small event, the splitting of the nucleus of an atom by one proton, rapidly transforms a radioactive system from matter to energy. There are two essential features of critically organized systems: 1) large-scale perturbations can be generated from very small events and 2) the organizational structure of the system and its elements render it sensitive to small changes. Two examples may be suggested:

1. A large group of dominoes is placed on a board. They are made to stand upright in a straight line so that the distance between one domino and its neighbor(s) is never greater than one domino length. A large bowling ball is dropped on the surface supporting the dominoes, slightly to the right of the beginning of the structure.

Result: all fall down.

2. The dominoes are re-erected so that they have the same configuration as described above. In this instance, however, no bowling ball is dropped. Rather, the first domino in the series is given a light tap that is just strong enough to topple it over onto its neighbor.

Result: all fall down.

A number of features of these examples deserve highlighting. First, one event (the dropping of the ball or delivering the light tap) precipitates a large change in the structure of each system. Second, although the two events have the same outcome, the precipitating events differ by orders of magnitude. The chain reaction described in the second example occurs because the distance between each element of this system permitted a small initiating event to grow to a large terminal event. Much like the atomic structure of radioactive material, the dominoes in this simple system are densely packed so that a small perturbation (as well as a large one) can produce catastrophic results. It should be noted that a dose-response model does not account for observed interactions among elements of this system because the energy expended in an initial tap can be orders of magnitude smaller than the energy expended in the collapse of the structure.

One can easily vary some of the conditions of this system by setting up the dominoes into clumps or clusters instead of lining them up in a row. By varying the distance between each cluster (and by varying the number of dominoes in each cluster), one is able, with just one tap, to create domino topplings of various sizes and shapes. Clusters that are closely grouped to one another will interact; clusters that are set at distances in excess of one domino length will not. The potential for creating a catastrophic alteration of the shape of each cluster, in other words, can be made to vary by changing the spacing of the elements in this system.[11]

As these examples suggest, a full understanding of some system outcomes depends more on the organization of the system than it does on the nature of each interacting element. Thus, the size of each domino is less important than the way in which all of the dominoes are configured. The degree to which such a system is critically organized is a function of the likelihood that large numbers of its component elements are sensitive to small changes in energy state.

Bak and colleagues[11,12] describe a paradigm for self-organized criticality in another simple system: a sandpile. An hourglass can be created that drops one grain of sand at measured intervals of time onto a plate some centimeters below. As the grains collect, a pile is created. After the passage of time, the slope of this pile will be such that an avalanche is precipitated with each new additional grain. This avalanche will vary in size, from a few grains to a large number of grains. The size varies both as a function of the amount of sand that has already fallen and as a function of the pile's previous history of large and small avalanches.

Sandpiles of this sort are an example of a critically organized system. At some point in the history of the pile, the number of grains falling off is roughly equal to the number being added in a given period of time. More precisely, the slope of the pile is such that each added grain of sand may precipitate: 1) no appreciable change in the configuration of the pile, 2) a small avalanche (involving a small number of grains), 3) a large avalanche (involving a larger number of grains), or 4) a "catastrophic" event in which a significant proportion of all of the grains will change position. As Bak and Chen observe:

> The sandpile has two seemingly incongruous features: the system is unstable in many different locations; nevertheless, the critical state is absolutely robust. On the one hand, the specific features, such as the local configurations of sand, change all the time because of the avalanches. On the other, the statistical properties, such as the size distribution of avalanches, remain essentially the same.[11]

Although the key events in this system are easily specified–falling grains of sand produce changes in a pile of grains–our ability to predict the occurrence of a large avalanche becomes less accurate the further into the future our predictions are made. We know that small avalanches are more likely to occur, for example, than large avalanches; we know that very, very large avalanches are relatively unlikely. We know that both classes of events will occur; we are unable, however, to predict when.

Although this inability to make accurate predictions may make such avalanches appear to be random events, the authors point out that they are not. A graph of sand movements plotted over time would show tremendous variation, but it also would display the characteristics of a specific kind of signal known as "flicker noise," so named because it belongs to the family of power functions, 1/f, or "one over 'ef' noise." Concretely, this means that the chance of a major avalanche is inversely proportional to the frequency of small avalanches. This "signal" is generated because the current behavior of a particular system is partially a function of its history, whereas "white noise" (a random signal) implies that current and past events are not correlated.

Critical systems, despite the seeming unpredictability, complexity, and apparent randomness of their behavior, arise from comparatively simple processes. The apparent randomness of the sandpile avalanches, for example, is actually, if one may entertain a counterintuitive notion, a deterministic randomness. The dynamics of the system are determined in the sense that the pile grows as the result of a simple, invariant process: the addition of one grain of sand at a measured interval. The system is also deterministic in the sense that avalanches are inevitable. What is not predetermined or predictable is the exact point in time when the sandpile will go from one metastable state to an avalanche, to another metastable state, and it is the uncertainty associated with our predictions that provide critical systems with their random flavor.

Bak and colleagues suggest that many systems in nature exhibit this kind of self-organized criticality. For example, stock markets periodically reach a critical state and crash. Similarly, the geologic system of tectonic plates in a region such as California will periodically undergo a shift in infrastructure that results in both major and minor earthquakes. As is the case with our sandpile model, the structure of the stock market and of the California tectonic plates displays this combination of inevitability and uncertainty–the inevitability of sandpiles, quakes, and crashes coupled with tremendous uncertainty as to when such events will occur.

The flow of patients through King/Drew Medical Center, or through any public hospital serving a poor, economically distressed community, appears to have many features of a critically organized system. This hospital has frequent eruptions of "emergency department gridlock"–that state of affairs in which "the configuration of interlocking stretchers" is so tight that it becomes difficult to walk from one nursing station to another. Such gridlock is an example of the avalanche-like conditions that are created when the demand for medical services exceeds the facility's carrying capacity. The tendency of this public health system to become gridlocked also suggests numerous scenarios in which "feedback" from a closed emergency department might precipitate the collapse of the general health of the community.

Multidrug-resistant tuberculosis, for example, provides the principal dynamic in one such scenario.

> A homeless intravenous drug user with tuberculosis enters the emergency department as the result of an asthma attack. Although he has been diagnosed and treated for his tuberculosis, he has not taken his medication for more than a month. He has an unusually virulent strain of multidrug-resistant tuberculosis. He remains for many hours in a poorly ventilated waiting room that is filled to overflowing. He coughs frequently. The length of his wait assures that many people are exposed. During the contagious but asymptomatic incubation period, many of these exposed individuals pass the strain to others, leading to an outbreak of epidemic proportions.

Wallace et al[13] provide a real world complement to this hypothetical example. They describe the "hollowing out" of the South Bronx between 1970 and 1980 when a series of catastrophic fires caused the loss of between 55% and 81% of occupied housing units within the Bronx's 62 contiguous Health Areas. They trace this firestorm to the disbanding or removal from service of some 50 New York City firefighting units. There was an additional cut of 20% to 25% of the number of firefighters in existing fire companies. As the authors note:

> From 1972 to 1976, following the fire service reductions, engine company structural fire worktime in New York City–a composite of building fire number and seriousness–rose from 44,000 to 63,000 hours, some 45%. The greatest increase was concentrated in precisely those areas which already had high fire rates, such as Brownsville, East New York, and the South Bronx; this accounts for the present "bombed out" aspect of many of these communities.[13]

The process described here unleashed a pattern that the authors describe as "contagious urban decay" in which a combination of fires and building abandonment drives the population of poor communities into increasingly overcrowded housing in the inhabitable units that remain. This hyperconcentration of people and poverty is associated with a concomitant collapse in the health of community members, specifically with respect to the spread of infectious diseases such as tuberculosis and increases in sociopathological behaviors such as violence and drug use. The authors also provide evidence that the disruption of social networks caused by the destruction of certain neighborhoods may have also moved clusters of intravenous drug users with HIV infections from neighborhood to neighborhood, effectively "seeding" the acquired immunodeficiency syndrome (AIDS) epidemic in neighboring communities where they continually formed and reformed networks of needle-sharing users.

Critically organized public health systems are not inevitably going to suffer catastrophic collapse. Careful examination of the elements of some critical systems suggests that they may be stabilized or coerced into a state of subcritical metastability. Wallace and colleagues,[13] for example, suggest that the instability created in communities that suffer high levels of overcrowding and dense concentrations of uncollected trash can be made resistant to contagious fire destruction by increasing the capacity of key blocks to resist catastrophic fires. Increased fire services and increased trash pickups in such settings would reduce the probability of fires in these blocks and, accordingly, increase the likelihood that the surrounding community will remain socially and structurally intact.

Ditto and Pecora,[14] for example, suggest that because complex systems exhibiting chaotic behavior often reflect the existence of many subsystems that are "out of phase," it should be possible to find the means to coax these systems back into a more stable state. In the examples that have been cited here, the provision of routine services such as regular trash pickups or the addition of one firefighter to the local firehouse might begin a process that would reverse the pattern of collapse that Wallace and colleagues describe.[13,15]

What are the features of solutions that might create such stability? The first requires an identification of those nodes within a critical system that generate the greatest levels of instability. The principle here is simple: in an unstable system composed of many interconnected subunits, the units that should be stabilized are those that have the greatest number of connections to other units of the system. Because these unstable units have the capacity to precipitate unstable behavior in their neighbors, it follows that the greater the stability of these core units, the greater the level of stability of the entire system.

The second point is that systems that exhibit self-organized criticality are dynamic and, therefore, constantly evolving. As a consequence, the solutions we develop to promote stability must be malleable and able to change as the system evolves. In our tuberculosis scenario, the man infected with tuberculosis is, indeed, at the origin of an expanding epidemic in his community, but the community system that provided a connection to the larger community was the emergency room. Stabilizing the emergency department might involve any number of alternative strategies: more efficient primary care services (so that a tuberculosis infection does not evolve into multidrug-resistant tuberculosis); many small clinics in different neighborhoods within the community instead of having one large, centrally located, emergency room; or effective screening programs at job sites, housing projects, schools, etc.

Poverty has many public health sequelae. We have suggested here that one of them is a hyperconcentration of poor people in urban areas where, eventually, private hospitals are bankrupted or where primary health services cannot be provided in a manner that is either efficient or cost effective. As private hospitals close and primary health-care services are shut down, the health-care system implodes. The emergency room in a public hospital becomes overloaded with cases that have nowhere else to go, and conditions are created for an eventual explosion. This kind of system becomes critically organized because only one point of entry is provided for those seeking care, and any dysfunction at this one point can ripple–as our examples have suggested–throughout the system. Similarly, providing other points of entry to health-care services can have the opposite effect. These alternative routes to care (or to screening) relieve the pressure and provide the potential for flexible responses to community health crises before they "go critical."

SUMMARY

Self-organized criticality offers more than a descriptive model or a doomsday forecast. We have tried to suggest that it is a paradigm for understanding the interconnections between apparently complex processes. At best, it suggests a method for finding the pressure points that can be used to bring unstable systems of public health services into greater levels of stability. The model enjoins us to understand that our goal is not to achieve equilibrium–that perfect match between the demand for health services and its delivery–but rather stability (or, more precisely, metastability). As is true of the sandpile, our systems of public health are constantly evolving. If we are correct, then the mechanism driving this ostensibly complex pattern of change and growth reflects the existence of simpler and, hopefully, more manageable processes. By monitoring these processes, it may be increasingly possible to adapt to change and even manage it effectively.

Literature Cited

1. Angell M. Privilege and health–what is the connection? *N Engl J Med.* 1993;329:126-127.
2. Shoemaker WC, James B, King LM, Hardin E, Ordog GJ. Urban violence in Los Angeles in the aftermath of the riots. *JAMA.* 1993;270:2833-2837.
3. Snow J. *On the Mode of Communication of Cholera.* 2nd ed. Reprinted as: *Snow on Cholera.* New York, NY: The Commonwealth Fund; 1956.
4. Traub JF, Wozniakowski H. Breaking intractability. *Sci Am.* 1994;270:102-107.
5. Lorenz EN. Cited in: Gleik J. *Chaos: Making a New Science.* London, England: Cardinal Press; 1987:20.
6. Snow J. *On the Mode of Communication of Cholera.* 2nd ed. Reprinted as *Snow on Cholera.* New York, NY: The Commonwealth Fund; 1956..
7. Sornette A, Sornette D. Self-organized criticality and earthquakes. *Europhysics Letters.* 1989;9:197-202.
8. Allen PM, Sanglier M. Urban evolution, self-organization, and decisonmaking. *Environment and Planning A.* 1981;13:167-183.
9. Schaffer WM, Kot M. Nearly one dimensional dynamics in an epidemic. *J Theor Biol.* 1985;112:403-427.
10. Prigogine I, Stengers I. *Order Out of Chaos: Man's New Dialogue With Nature.* London, England: Flamingo; 1984.
11. Bak P, Chen K. Self-organized criticality. *Sci Am.* 1991;264:46-53.
12. Bak P, Tang C, Wiesenfeld K. Self-organized criticality. *Physical Review A.* 1988;38:364-374.
13. Wallace R, Fullilove MT, Wallace D. Family systems and deurbanization: implications for substance abuse. In: Lowinson JH, Ruiz P, Millman R, eds. *Substance Abuse: A Comprehensive Textbook.* 2nd ed. Baltimore, Md: Williams and Wilkins; 1992:944-955.
14. Ditto WL, Pecora LM. Mastering chaos. *Sci Am.* 1993;269:78-84.
15. Wallace R, Wallace D. The coming crisis of public health in the suburbs. *Milbank Q.* 1993;71:543-564.