A simple guide to chaos and complexity
Dean Rickles, Penelope Hawe, Alan Shiell

The concepts of complexity and chaos are being invoked with increasing frequency in the health sciences literature. However, the concepts underpinning these concepts are foreign to many health scientists and there is some looseness in how they have been translated from their origins in mathematics and physics, which is leading to confusion and error in their application. Nonetheless, used carefully, “complexity science” has the potential to invigorate many areas of health science and may lead to important practical outcomes; but if it is to do so, we need the discipline that comes from a proper and responsible usage of its concepts. Hopefully, this glossary will go some way towards achieving that objective.

Glossary

The concepts of complexity and chaos are being increasingly invoked in the health sciences literature (general treatments are outlined in references1–4). Applications to date include (there are many more):

- epidemiology and infectious disease processes5–14
- healthcare organisation15–26
- general practice and “the clinical encounter”27–30
- biomedicine31–54
- health social science35–56
- health geography37,38

However, despite their being so widespread, the concepts underpinning complex systems science and chaos theory are still foreign to many health scientists and there is some looseness in how they have been translated from their origins in mathematics and physics, which is leading to much confusion and error in their application.4 This glossary attempts to resolve these issues by providing a simple (but not simplified) guide to many central concepts in chaos theory and complexity science. The many references provide more detail.

Dynamical systems

System is simply the name given to an object studied in some field and might be abstract or concrete; elementary or composite; linear or nonlinear; simple or complicated; complex or chaotic. Complex systems are highly composite ones, built up from very large numbers of mutually interacting subunits (that are often composites themselves) whose repeated interactions result in rich, collective behaviour that feeds back into the behaviour of the individual parts. Chaotic systems can have very few interacting subunits, but they interact in such a way as to produce very intricate dynamics. Simple systems have very few parts that behave according to very simple laws. Complicated systems can have very many parts too, but they play specific functional roles and are guided by very simple rules. Complex systems can survive the removal of parts by adapting to the change: to be robust, other systems must build redundancy into the system (eg by containing multiple copies of a part). A large healthcare system will be robust to the removal of a single nurse because the rest of the members of the system will adapt to compensate—however, adding more nurses does not necessarily make the system more efficient.21–24,37

On the other hand, a complicated piece of medical technology, such as a positron emission tomography scanner, will obviously not survive removal of a major component. The behaviour of a chaotic system appears random, but is generated by simple, non-random, deterministic processes: the complexity is in the dynamical evolution (the way the system changes over time driven by numerous iterations of some very simple rule), rather than the system itself.39–46

Systems possess properties that are represented by variables or observables: quantities that have a range of possible values such as the number of people in a population, the blood pressure of an individual, the length of time between consultations, and so on. The values taken by a system’s variables at an instant of time describe the system’s state. A state is often represented by a point in a geometrical space (phase space), with axes corresponding to the variables. The coordinates then correspond to particular assignments of values to each variable. The number of variables defines the dimension of both the space and the system. Each point in the phase space represents a way in which the system could be at a time, corresponding to an assignment of particular values to the variables at an instant. The overall health state of an individual, for example, might include values for lung capacity, heart rate, blood-sugar levels and so on. A path through the phase space corresponds to a trajectory of the system, or a way in which the system could evolve over time—for example, the change in a person’s health state over time (itself a function of many other variables).

A dynamical system is a system whose state (and variables) evolve over time, doing so according to some rule. How a system evolves over time depends both on this rule and on its initial conditions—that is, the system’s state at some initial time. Feeding this initial state into the rules generates a solution (a trajectory through phase space), which explains how the system will change over time; chaos is generated by feeding solutions...
Complex and chaotic systems are both examples of nonlinear dynamical systems. A linear system is characterised by the satisfaction of the superposition principle. The superposition principle says that if A and B are both solutions for some system (ways in which the system could evolve), then so is their sum A + B—this implies that a linear system can be decomposed into its parts and each part solved separately to construct the full solution. For nonlinear systems, this is not possible because of the appearance of nonlinear terms, functions of the variables such as sin(x), x^n and xy. In this sense then, the whole here is more than the sum of its parts. Given this, a nonlinear system is one for which inputs are not proportional to outputs: a small (large) change in some variable or family of variables will not necessarily result in a small (large) change in the system. This kind of behaviour is well known to those involved in intervention research: large interventions, in some variable, do not necessarily have a large effect on some outcome variable of interest. Likewise, a small intervention can have large, unexpected outcomes.

A system (or process) is deterministic if it is possible to uniquely determine its past and future trajectories (ie all points it passed through and will pass through) from its initial (present) state. A system is semi-deterministic if its future but not its past trajectory can be uniquely determined. An indeterministic system is one without a unique future trajectory, so that the evolution is random. It is often possible to tell whether or not a system is deterministic by inspecting the time series it generates, plotting states at different times. The data points in the time series might correspond to measurements of blood sugar levels, population health indicators and so on. If closely matched points are found in different places in the series, then the series is investigated to see if the points that follow are closely matched too. Hence, attempts are made to infer from the spread of points the kind of system or process that generated it.

The time-series can also be used to indicate whether or not a system is chaotic or complex: a chaotic system’s time-series has a fractal-like structure, meaning that it looks the same at different scales (a property also known as self-similarity: take a snapshot of the time series covering a certain interval of time and then take another snapshot covering a much larger period of time and the two will look the same). Fractals are really just a spatial version of chaos: the interesting type of chaos is the temporal kind. Inferring complexity from a time-series is more difficult but can be done (this involves finding power law correlations in the data; see below).

A related concept is that of the attractor. Following an intervention in a system (changing the value of some variable), it takes a little while for it to settle down into its “normal” behaviour. The path traced out in phase space during this period is known as a transient (or the start-up transient). The state it reaches after this corresponds to the normal behaviour of the system: the phase space points corresponding to this form the system’s attractor. If the attractor is a point that does not move, it is known as a fixed point. Such attractors often describe dissipative systems (those that lose energy—for example, due to friction). In general, however, it is possible to think of an attractor as whatever the system behaves like after it has passed the transient stage.

There are other types of attractor. For example, an attractor that describes a system that cycles periodically over the same set of states, never coming to rest, is known as a limit cycle. A system need not possess a single attractor either; the phase space can have a number of attractors whose “attractiveness” depends upon the initial conditions of the system (ie the state of the system at the outset).

The set of points that are “pulled” towards a particular attractor are known as the basin of attraction. What is important to note about the types of attractor discussed above (fixed points and limit cycles) is that initial points that are close to each other remain close as they each evolve according to the same rules. This property is violated for so-called strange attractors for which close points diverge exponentially over time. Such attractors are also aperiodic; systems described by strange attractors do not wind up in a steady state nor do they repeat the same pattern of behaviour (an example that can be easily observed in the household is a dripping water tap tuned to a certain flow, not too high and not too low). Chaotic systems have strange attractors; complex systems have evolving phase spaces and a range of possible attractors.

These concepts have been applied extensively, accurately and successfully in the biomedical sciences. The general outcome of these investigations appears to be that chaos is associated with “good health”: pathologies (such as of the brain, heart, lungs) occur when the dynamics becomes stable and the attractor is a limit cycle. Heart disease, epilepsy, bipolarity and so on are considered to be dynamic diseases in that they are not associated with something that can manifest itself at an instant (like a broken arm), but only arise over time. This is why chaos is so appropriate.

**Chaos versus complexity**

We can now consider further the similarities and differences between chaotic systems and complex systems. Each shares common features, but the two concepts are very different. Chaos is the generation of complicated, aperiodic, seemingly random behaviour from the iteration of a simple rule. This complicatedness is not complex in the sense of complex systems science, but rather it is chaotic in a very precise mathematical sense. Complexity is the generation of rich, collective dynamical behaviour from simple interactions between large numbers of subunits. Chaotic systems are not necessarily complex, and complex systems are not necessarily chaotic (although they can be for some values of the variables or control parameter; see below).

The interactions between the subunits of a complex system determine (or generate) properties in the unit system that cannot be reduced to the subunits (and that cannot be readily deduced from the subunits and their interactions). Such properties are known as emergent properties. In this way it is possible to have an upward (or generative) hierarchy of such levels, in which one level of organisation determines the level above it, and that level then determines the features of the level above it. Emergent properties may also be universal or multiply realisable in the sense that there are many diverse ways in which the same emergent property can be generated. For example, temperature is multiply realisable: many configurations of the same substance can generate the same temperature, and many different types of substance can generate the same temperature.

The properties of a complex system are multiply realisable since they satisfy universal laws—that is, they have universal properties that are independent of the microscopic details of the system. Emergent properties are neither identical with nor reducible to the lower-level properties of the subunits because there are many ways for emergent properties to be produced. A necessary condition, owing to nonlinearity, of both chaos and complexity is sensitivity to initial conditions. This means that two states that are very close together initially and that operate under the same simple rules will nevertheless follow very different trajectories over time. This sensitivity makes it difficult to predict the evolution of a system, as this requires

www.jech.com

Rickles, Hawe, Shiel
the initial state of the system to be described with perfect accuracy. There will always be some error in how this is performed and it is this error that gets exponentially worse over time. It is possible to see how this might pose problems for replication of initial conditions in various types of trial and intervention.24 25

There are several less well-understood, but nonetheless important properties that are characteristic features of complex systems. Complex systems often exhibit self-organisation, which happens when systems spontaneously order themselves (generally in an optimal or more stable way) without “external” tuning of a control parameter (see below). This feature is not found in chaotic systems and is often called anti-chaos.26 Such systems also tend to be out of equilibrium, which means that the system never settles in to a steady state of behaviour. This is related to the concept of openness: a system is open if it is not or cannot be screened off from its environment. In closed systems, outside influences (exogenous variables) can be ignored. For open systems, this is not the case. Most real-world systems are open, thus this presents problems both for modelling and experimenting on such systems, because the effect of exogenous influences must be taken into account. Such influences can be magnified over time by sensitivity to initial conditions.

Another important feature of a complex system is the idea of feedback, in which the output of some process within the system is “recycled” and becomes a new input for the system. Feedback can be positive or negative: negative feedback works by reversing the direction of change of some variable; positive feedback increases the rate of change of the variable in a certain direction. In complex systems, feedback occurs between levels of organisation, micro and macro, so that the micro-level interactions between the subunits generate some pattern in the macro-level that then “back-reacts” onto the subunits, causing them to generate a new pattern, which back-reacts again and so on. This kind of “global to local” positive feedback is called coevolution, a term originating in evolutionary biology to describe the way organisms create their environment and are in turn moulded by that environment.

If a system is stable under small changes in its variables, so that it does not change radically when interventions occur then it is said to be robust. Generally, complex systems increase in robustness over time because of their ability to organise themselves relative to their environment. However, it is possible for single events to alter a complex system in a way that persists for a long time (this is called path-dependence). For a complex system, “history matters”. Complex processes (processes generated by complex systems) are, therefore, non-Markovian: they have a long “memory”. This non-Markovian aspect (essentially due to feedback mechanisms) is often believed to pose insuperable problems for describing causal events and making causal predictions about complex systems. If this were so, it would be a severe blow to the practical usefulness of complex science; however, there has been some encouraging preliminary work involving non-Markovian and modified Markovian causal models and networks.63 64

The key differences between chaotic systems and complex ones lie, therefore, in the number of interacting parts and the effect that this has on the properties and behaviour of the system as a whole. As Nobel Laureate Phillip Anderson put it: “more is different”.65 Complex systems are coherent units in a way that chaotic systems are not, involving instead interactions between units. This simple difference concerning units and subunits can be brought out using concepts from the theory of critical phenomena, which is central to complexity science.

Critical phenomena
Some systems have a property known as criticality. A system is critical if its state changes dramatically given some small input. To make sense of this, we need to introduce several new concepts.

An order parameter is a macroscopic (global or systemic) feature of the system that tells one how the parts of the system are competing or cooperating with one another. Hence, there is a state of order among the parts when they act collectively (in which case the order parameter is non-zero) and a state of disorder when they fight against each other, doing the opposite of their neighbours (in which case the order parameter is zero).

The control parameter is an external input to the system that can be varied so as to change the order parameter and so the macroscopic features of the system. This is termed tuning the control parameter to shift the system between various phases or regimes; it is possible to have ordered, chaotic and critical (edge of chaos) phases. A common example concerns the phenomenon of magnetism in a piece of metal. Here the order parameter is the degree of magnetism and the control parameter is temperature. As the magnet is heated up, the magnetism decreases; increasing the control parameter decreases the order parameter. The system undergoes a phase transition so that at a critical temperature the magnetism vanishes. This is a general feature of complex systems; tuning the system’s control parameter to a certain critical point results in a phase transition at which the system undergoes an instantaneous radical change in its qualitative features (the phases of water, from gas to liquid to solid, is another common example). The general study of such behaviour is the theory of critical phenomena.

A system that is at a critical point has an extremely high degree of connectivity between its subunits: everything depends on everything else! Complex systems are said to be poised at such a position, between order and chaos. The degree of connectedness is encoded in the correlation function. This function tells us how much pairs of subunits are influencing one another and how much this influence varies with distance. The furthest distance the influence extends is known as the correlation length; beyond this distance, the subunits are independent and are unaffected by one another. The farther apart two subunits are, the less they influence each other (the influence effect decreases exponentially with distance). The correlation length can itself be influenced by the control parameter. When the control parameter is tuned to the critical point, the correlation length becomes infinite, and all the subunits follow each other. The influence still decays exponentially with distance, but in the critical regime, more pathways are opened up between pairs of subunits so that the connectivity of the system is massively amplified. As a consequence, a small disturbance in the system (even to a single subunit) can produce massive, systemic changes. Very different kinds of critical system exhibit the same properties—for example, crowds behaving like the atoms in a magnet. This feature is known as universality.14

Universality is connected to another concept: scaling. Scaling laws, or power laws have the following form: f(x) ∝ x−α in other words, the probability f(x) of an event of magnitude x occurring is inversely proportional to x.12 Sociologist Vilfred Pareto noticed that the statistical distribution for the wealth of individuals in a population followed such a law: few are rich, some are well-off and many are poor. Whenever systems are described by scaling laws that share the same exponents (the α term) then they will exhibit similar behaviour in some way (hence, universality). Such systems are said to belong to the same universality class.16 17 For example, diverse countries follow Pareto’s law, as do cities. This result could easily be extrapolated to the distribution of the health of individuals within and across populations, with significant implications for research on health inequalities.

If a system displays power law behaviour then it is scale-free and its parts have scale-invariant correlations between them. What
What this paper adds

This glossary is intended to provide a “corrective” to what the authors perceive to be a loose and misleading handling of ideas stemming from the sciences of complexity. The aim is to provide as simple an account as possible, without succumbing to the “popular science”-style accounts that appear to be fast becoming the norm in the health-sciences literature. The going is perhaps a little tough, but we feel the payoff (a responsible guide) is worth the extra effort.

this means is that the system does not possess a characteristic scale in the sense that events of all magnitudes can occur. To make sense of this, take an example of a system that does have a characteristic scale: human height. Most humans are of about the same height. There are a few “outliers” who are either very tall or very small, but most are around 5’8”. Now consider earthquakes. These can be both extremely tiny and extremely massive, but most are negligible. This is an example of a system that does not have a characteristic scale; it satisfies a scaling law (the Gutenberg-Richter law). Note that self-similarity, scale invariance and power laws are just equivalent ways of saying that there is no characteristic scale. Various empirical studies have confirmed that healthcare systems obey power laws for a number of quantities of interest, such as hospital waiting lists. These studies have clear policy implications; if various healthcare delivery systems exhibit power-law behaviour then we ought to reframe our intervention and management theories in terms of complex systems science. We should not be surprised if huge catastrophes occur for no discernible reason, and we should not be surprised if our massive intervention to reduce waiting times by employing more staff does nothing for efficiency or even makes it worse. However, we should show caution in using the power-law behaviour to infer an underlying complex system, as power laws can be generated in diverse ways.

CONCLUSION

Although complexity science is still a work in progress, having neither a firm mathematical foundation nor the necessary and sufficient conditions whose satisfaction entails that some system is complex, it nonetheless has the potential to invigorate many areas of health research as we hope to have indicated in this brief guide. However, if it is to achieve its potential we need the discipline that comes from careful and proper usage of its concepts. Hopefully, this glossary will go some way towards achieving that objective.

Authors’ affiliations

Dean Rickles, Department of Philosophy, University of Calgary, Calgary, Canada
Penelope Hawe, Alan Shiel, Markin Institute, University of Calgary, Calgary, Canada

Funding: This work was conducted as part of an International Collaboration on Complex Interventions (ICCI) funded by the Canadian Institutes of Health Research. DR is an ICCI Post Doctoral Fellow. AS and PH are Senior Scholars of the Alberta Heritage Foundation for Medical Research. PH holds the Markin Chair in Health and Society at the University of Calgary.

REFERENCES

35 Kernick D. Wanted—new methodology for health services research. Is complexity theory the answer? Fam Pract 2006;23:385–90.

www.jech.com