Deconstructing HARI: On The Dynamical Analysis of Clinical Trial Outcomes

George A. Diamond

Division of Cardiology, Cedars-Sinai Medical Center, and the David Geffen School of Medicine, University of California, Los Angeles, California, USA

Abstract

Stochastic survival models, such as Cox proportional hazards regression, are commonly used to assess outcomes in randomized clinical trials. These models are typically designed as non-deterministic two-state systems representing the transition from an initial baseline state to a terminal outcome state. Many pathophysiological processes, however, involve a deterministic sequence of additional transitions: from the baseline state to an intermediate state (e.g., unstable angina) and thence to a terminal state (e.g., myocardial infarction). In such cases, the typical two-state model may be insufficient to capture the dynamics of the actual clinical process. The purpose of this essay is to compare the performance of a three-state dynamical model to a two-state stochastic model in a hypothetical clinical trial. This comparison demonstrates that the former is better than the latter (i) at deconstructing the sequence of outcome transitions into its individual causal components, and (ii) at correctly characterizing the effect of treatment on each of these components. In conclusion, dynamical models can provide novel insights into the mechanism(s) by which a given treatment is associated with benefit and harm. Such models should therefore be used more widely in the design, analysis and interpretation of clinical trials.

Keywords: Biostatistics; Clinical trials; Mathematical models

Deterministic dynamical models versus non-deterministic stochastic models

The distinction between dynamical and stochastic models is likely to be unfamiliar to clinicians and statisticians alike. In brief, the former describes the behavior of causal deterministic systems (the usual domain of physics), while the latter describes the behavior of probabilistic non-deterministic systems (the usual domain of epidemiology).

Dynamical models have their origins in late 17th Century Newtonian mechanics, as a way of describing the deterministic behavior of complex systems such as planetary orbits in terms of the solutions to ordinary differential equations. In 1864, Waage and Guldberg expanded dynamical modeling to the field of chemistry by formulating the “law of mass action,” which states that the speed of a chemical reaction is proportional to the quantity of the reacting substances [9]. The resultant discipline of chemical kinetics thereby describes how different conditions can influence the speed of reaction and yield information about its mechanism [10].

Stochastic models have their origins in late 19th Century applications of probability as a way of describing the non-deterministic behavior of complex systems (such as economics), using sets of random variables to predict the evolution of random events over time. In contrast to a dynamical model which can evolve only in a single way (as does the solution to an ordinary differential equation), a stochastic model can evolve in many ways.

Received August 13, 2015; Accepted September 30, 2015; Published October 05, 2015


Copyright: © 2015 Diamond GA. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.
The dynamical/stochastic distinction is the basis for the epistemological divide between classical and Bayesian statistics (the former considering hypotheses to be fixed and observations to be random variables, and the latter considering observations to be fixed and hypotheses to be random variables). As we shall see, this seemingly academic distinction between the realism of empirical observation and the idealism of theoretical interpretation has important clinical consequences.

The initial application of dynamical models to medicine dates back to the opening paragraph of a 1926 paper by A. G. McKendrick [11], a pioneer of mathematical epidemiology:

In the majority of the processes with which one is concerned in the study of the medical sciences, one has to deal with assemblages of individuals, be they living or be they dead, which become affected according to some characteristic... If one thinks of these individuals...as moving in all sorts of dimensions, reversibly or irreversibly, continuously or discontinuously, by unit stages or per saltum, then the method of their movement becomes a study in kinetics, and can be approached by the methods ordinarily adopted in the study of such systems [12].

McKendrick’s principal focus was on the analysis of epidemics and infectious disease transmission. Since then, his dynamical approach has been applied to a broad range of biological processes including plasma membrane electrodynamics [13], digitalis pharmacokinetics [14], contractile periodicity [15], myocardial imaging [16], cancer chemotherapy [17], and clinical epidemiology [18-21].

The two-state model

In its simplest instantiation, a two-state dynamical model \((A \rightarrow B)\) can be used to represent the transition from an initial state \(A\) (analogous to the substrate of a chemical reaction) to a subsequent state \(B\) (analogous to the product of a chemical reaction). Mathematically, this model can be represented by an ordinary differential equation, where the rate of change in \(A\) with respect to time, \(t\), is inversely proportional to its prevalence and to a constant of proportionality or rate constant, \(k\) (the hazard), which can be quantified by fitting empirical data on \([A]\) and its rate of change to the dynamical rate equation:

\[
\frac{d[A]}{dt} = -k[A]
\]

As shown in the next paragraph, the precision of this parameterization is proportional to the sample size employed in its estimation.

Let us compare this two-state dynamical model to a two-state stochastic model in a hypothetical clinical trial. The control group consists of 1000 patients with UA (defined by some set of entry criteria), 500 of whom manifest MI over 90 days of follow-up; the treatment group consists of an additional 1000 patients (defined by the same set of entry criteria), only 250 of whom manifest MI over 90 days of follow-up. Based on these empirical observations, and the aforementioned equations for a two-state dynamical model, \(k = 0.0077 \pm 0.0004\) days\(^{-1}\) (\(-\ln(1-500/1000)/90\)) for the control group vs \(0.0032 \pm 0.0002\) days\(^{-1}\) (\(-\ln(1-250/1000)/90\)) for the treatment group. Using these data, a random sample of 1000 patients for each group was generated in which the rate constants are log-normally distributed.

\[\text{[Para 107]}\]

\[\text{[Para 108]}\]

and time-to-event is exponentially distributed. Conventional survival analyses were performed on these patient level data using WinSTAT (version 2012.1). The resultant stochastic and dynamical survival curves were almost identical (Figure 1), and the stochastic hazard ratio of 0.45 (95% CI: 0.39-0.52; p<0.00001) based on Cox regression was very similar to the dynamical hazard ratio of 0.42 (95% CI: 0.36-0.48; p<0.00001) based on the ratio of the rate constants. Using a simple two-state model, therefore, the dynamical analysis and stochastic analyses are essentially equivalent.

The three-state model

In contrast, a three-state dynamical model \((\dot{A} \rightarrow B \rightarrow C)\) represents the state transitions, not as a single first order process, but as a sequence of first order processes \(\dot{A} \rightarrow B\) and \(\dot{B} \rightarrow C\). This three-state model can be used to mirror a wide range of clinically relevant scenarios, from \(k_1 < k_2\) to \(k_1 \approx k_2\) to \(k_1 < k_2\) (Figure 2). Consider the scenario when \(k_1 > k_2\) (say, by a factor of 5). As a result, the transition \(A \rightarrow B\) will predominate before the transition \(B \rightarrow C\) gets underway. Consequently, the intermediate \([B]\) increases rapidly at a rate defined by \(k_1\), and thereafter decreases more slowly at a rate defined by \(k_2\). The situation has been likened to three vertically positioned water buckets [10]. Imagine the top bucket leaks through a large hole into the middle bucket, which leaks in turn through a much smaller hole into the bottom bucket. As a result, water placed in the first bucket flows rapidly into the second bucket, from which it then flows more slowly into the third bucket. This scenario is analogous to the relatively fast initial transition from UA to MI and to the relatively slow subsequent transition from MI to cardiac death (the lower left panel in Figure 2).

A contrasting scenario occurs when we place the small hole in the top bucket and the large hole in the middle bucket, thereby making \(k_1 < k_2\) (again, by a factor of 5). Now, water leaks out of the middle bucket faster than it comes in. As a result, the water level in the middle bucket never rises very high. Consequently, \([B]\) rapidly approaches a value approximated by the ratio \(k_3/k_1\), and then falls more slowly at a rate defined by \(k_2\). Because \(k_3/k_2\) is relatively low, \([B]\) is similarly low. This scenario is analogous to the relatively slow initial transition from stable coronary disease to stent thrombosis and the relatively fast subsequent transition from stent thrombosis to cardiac death (the upper right panel in Figure 2).

Even more complex dynamical models have been constructed to assess independent outcomes representative of safety and efficacy [19] competitive outcomes exemplified by cardiovascular versus
The empirical observations over 90 days of follow-up (Table 2) demonstrated statistically significant reductions in this outcome (Table 2), but the Kaplan-Meier curves fail to capture the actual pattern of the change over time. Since $k_2 < k_3$, in this example, the dynamical model shows that [UA] remains relatively low and invariant over follow-up in both the control and treatment groups. This is consistent with our water bucket analogy. If you reduce the size of the hole in the first bucket, less water flows into the second bucket. The Kaplan-Meier curves derived from the very same data, however, exhibit a continual increase in [UA] over time for both the control and treatment groups, and do not therefore distinguish between cause (the change in $k_1$) and effect (the change in [UA]),

Effect of the initial transition rate on the terminal state

The pair of graphs in the upper right quadrant of Figure 3 illustrates the effect of the same 50% reduction in $k_1$ on the proportion of the terminal state [MI]. Now, both the stochastic and dynamical models exhibit similar patterns of reduction over follow-up (although the former overestimates the magnitude of this change by approximately 50% in comparison to the latter). Referring again to our water bucket analogy, since we reduced the size of the hole in the first bucket, causing less water to flow into the second bucket, less water also flows into the third bucket even though the hole in the second bucket has not changed in size. Consequently, changes in the initial transition rate echo through the sequence of subsequent transitions. The change in [MI] is thereby identified as statistically significant by both the stochastic and dynamical models even though there was no change in its transition rate (Table 2). The dynamical model identifies this disconnect between cause (the change in $k_1$) and effect (the change in [UA]); the stochastic model does not.

Effect of the subsequent transition rate on the intermediate state

The pair of graphs in the lower left quadrant of Figure 3 illustrates the effect of a 50% reduction in $k_2$, on the proportion of the intermediate state [UA]. The patterns of change were similar to those already described for $k_1$, but opposite in direction. In contrast to the decrease associated with a reduction in $k_1$, [UA] increased in response to a reduction in $k_2$, albeit not to a statistically significant degree (Table 2). This again is explained by our water bucket analogy. By reducing the size of the hole in the second bucket, less water flows into the third bucket, causing the

Figure 2: A sequential three-state dynamical model. The grey lines represent the intermediate state and the black lines represent the terminal state. Note the wide range of patterns for rate constants ($k_1$ and $k_2$) that vary over 1.4 orders of magnitude (from 0.002 to 0.05).
The graphs summarize the statistical analysis of a hypothetical randomized clinical trial. The top row illustrates the analyses based on a 50% reduction in \( k_1 \) (transition from the initial to the intermediate state). The bottom row illustrates parallel analyses based on a 50% reduction in \( k_2 \) (transition from the intermediate to the terminal state). The comparative analyses of these data are summarized in Table 2. See text for discussion.

### Table 1: Hypothetical Assessment of Revascularization Interventions (HARI)

<table>
<thead>
<tr>
<th>Transition</th>
<th>Control Group</th>
<th>Events</th>
<th>Pts</th>
<th>Days</th>
<th>( k \pm \text{SEM} )</th>
<th>Treatment Group</th>
<th>Events</th>
<th>Pts</th>
<th>Days</th>
<th>( k \pm \text{SEM} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( k_1 )</td>
<td>CAD→UA</td>
<td>165</td>
<td>1000</td>
<td>90</td>
<td>0.0020 ± 0.0002</td>
<td>( k_2 )</td>
<td>163</td>
<td>165</td>
<td>45</td>
<td>0.0981 ± 0.0156</td>
</tr>
<tr>
<td>( k_2 )</td>
<td>UA→MI</td>
<td>165</td>
<td>1000</td>
<td>90</td>
<td>0.0020 ± 0.0002</td>
<td>( k_2 )</td>
<td>163</td>
<td>165</td>
<td>45</td>
<td>0.0981 ± 0.0156</td>
</tr>
</tbody>
</table>

\( Pts=\)Patients, \( k=\)kinetic rate constant (\( \ln(1/\text{Events/Pts/\text{Days}}) \)), \( \text{SEM}=\)standard error of the mean, \( \text{CAD}=\)coronary artery disease, \( \text{UA}=\)unstable angina, \( \text{MI}=\)myocardial infarction, \( \text{CAD}→\text{UA}=\)transition from the initial state (CAD) to the intermediate state (UA), \( \text{UA}→\text{MI}=\)transition from the intermediate state (UA) to the terminal state (MI), \( \downarrow k_1 = \) reduced initial transition rate, \( \downarrow k_2 = \) reduced subsequent transition rate. The numbers in boldface represent 50% reductions in the post-treatment rate constants.

### Table 2: Stochastic and dynamical analysis of HARI

<table>
<thead>
<tr>
<th></th>
<th>Intermediate State</th>
<th></th>
<th>Terminal State</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Stochastic Model</td>
<td>Dynamical Model</td>
<td>Stochastic Model</td>
<td>Dynamical Model</td>
</tr>
<tr>
<td>( k_1 )</td>
<td>HR 0.32</td>
<td>0.50</td>
<td>HR 0.57</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>95% CI 0.17-0.61</td>
<td>0.38-0.65</td>
<td>0.43-0.75</td>
<td>0.59-1.73</td>
</tr>
<tr>
<td>( p )</td>
<td>0.001</td>
<td>0.000</td>
<td>0.000</td>
<td>0.973</td>
</tr>
<tr>
<td>( k_2 )</td>
<td>HR 1.30</td>
<td>1.00</td>
<td>HR 0.85</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>95% CI 0.87-1.93</td>
<td>0.81-1.24</td>
<td>0.67-1.09</td>
<td>0.35-0.73</td>
</tr>
<tr>
<td>( p )</td>
<td>0.198</td>
<td>1.000</td>
<td>0.208</td>
<td>0.000</td>
</tr>
</tbody>
</table>

HR= hazard ratio, CI= confidence interval, \( p=\)value, \( \downarrow k_1 = \) reduced initial transition rate, \( \downarrow k_2 = \) reduced subsequent transition rate. Values for the stochastic model are based on Cox regression; values for the dynamical model are based on a z test of the rate constants.

Effect of the subsequent transition rate on the terminal state

Finally, the pair of graphs in the lower right quadrant of Figure 3 illustrates the effect of the same 50% reduction in \( k_2 \) on \( [MI] \). The pattern of response was similar to that following a change in \( k_1 \), but less in magnitude. As before, reducing the size of the hole in the second bucket causes less water to flow into the third bucket, but this is now partially offset by the fact that reducing the size of the hole also causes the water level in the second bucket to rise, thereby increasing the flow of water into the third bucket. Consequently, despite a statistically significant reduction in \( k_2 \), the change in \( [MI] \) was not statistically significant (Table 2).

Clinical implications

Treatments can wittingly or unwittingly target a spectrum of pathophysiological mechanisms associated with efficacy and safety. A new drug, for example, might activate an anti-inflammatory pathway and stabilize the atherosclerotic plaque, thereby preventing the development of \( UA \), or it might inhibit a pathway involved in platelet aggregation and thereby prevent the occurrence of \( MI \).
distinctions can be masked, however, in the design of a clinical trial if the primary outcome of interest is taken to be a composite of causally interrelated events (UA or MI, for example). Although the components of such composites are usually reported along with the primary outcome, formal statistical analysis of these components can be hampered by limitations in sample size or baseline event rate [22].

In this context, a recent study compared the performance of alternative stochastic models for analysis of repeated ischemic events among all components of the primary endpoint (all cause death, MI or stroke) in the Targeted Platelet Inhibition to Clarify the Optimal Strategy to Medically Manage Acute Coronary Syndromes (TRILOGY ACS) trial (23). The investigators concluded that models accounting for all events, especially those incorporating subjective weightings indicative of the clinical relevance of the individual components, appeared most advantageous.

Thus, even when widely accepted methods of survival analysis are applied to idealized samples of data (as in HARI), we might yet be led to draw erroneous conclusions. Based on the Kaplan-Meier curves in Figure 3 and the Cox regressions in Table 2, for example, a drug that truly prevents UA (by halving the rate of plaque destabilization) may be falsely adjudged to prevent MI as well, when in fact the latter is only a passive consequence of the former. Similarly, a drug that truly prevents MI (by halving the rate of platelet aggregation) may be falsely adjudged to hasten the onset of UA (in this case, by 30%), and may not therefore be characterized as beneficial. A dynamical analysis circumvents these errors, and better identifies the underlying mechanism(s) of action.

Limitations

A few caveats are deserving of mention. First, the dynamical models in these analyses are not rigorously deterministic. A rigorously deterministic model would be able to tell us (no matter where we are at present) where we have been in the past and where we will be in the future. Each state in such a model would have a single unique arrow leading into it and a single unique arrow leading out of it. Moreover, a rigorously deterministic model would be reversible: change the direction of every arrow, and the model remains deterministic. Physics is like this (24); medicine is not (25). In medicine, we cannot usually reverse the sequence A → B → C to C ← B ← A. Even if we could, the model would still not be rigorously deterministic since there is no arrow telling us how we got to the initial state in the first place.

Second, the pedagogical design of HARI as a three-state consecutive process is an over simplification of a more complex network of transitions. In actuality, things are not so simple. Clinical events such as MI, for example, can occur more than once in a given patient. In addition, MI can occur without UA, and UA can resolve without MI. Although each of these complicating features can be modeled by increasingly complex sets of differential equations, and software exists to facilitate their solution, the task can be daunting.

Third, the dynamics of these models are assumed to be first order. In chemical kinetics, the order of an elementary transition is generally equivalent to its “molecularity” (the number of reactants involved in the process). By analogy, since clinical transitions typically involve individual patients (with the possible exception of communicable diseases and sexual contact), it is reasonable to consider them unimolecular processes manifesting linear first order dynamics [11]. If the empirical evidence suggests otherwise, the equations can again be modified to reflect this.

Fourth, both the stochastic and dynamical models are continuous analogs of an underlying discrete binomial process. Their accuracy therefore depends on the overall size of the study population and the attendant number of events in that population. The larger the sample, the better the models can be expected to perform.

Finally, while it would be preferable to base these analyses on patient level time-to-event data from an actual clinical trial, such data were not available in the medical literature. The analyses in this essay were therefore based on simulated data sets. Consequently, the conclusions deriving from these simulations will need to be verified by application to actual clinical trial data.

Conclusions

Notwithstanding these limitations, the present findings indicate that three-state dynamical models are better than two-state stochastic models (i) at deconstructing the sequence of outcome transitions into its individual causal components, and (ii) at correctly characterizing the effect of treatment on each of these components. Accordingly, these dynamical models can provide novel insights into the mechanism(s) by which a given treatment is associated with benefit and harm. Such models should therefore be used more widely in the design, analysis and interpretation of clinical trials.

Acknowledgments

This paper is an original composition of the author who is solely responsible for its content. There are no acknowledgements, funding sources or conflicts-of-interest to be disclosed.

References


