



Archimedes: a new model for simulating health care systems—the mathematical formulation

Leonard Schlessinger^a and David M. Eddy^{b,*}

^a Department of Biomathematical Analysis, Care Management Institute, Kaiser Permanente Program Office, Pasadena, CA, USA

^b Southern California Kaiser Permanente, 1426 Crystal Lake Road, Pasadena, CA, USA

Received 13 November 2001

Abstract

This paper designs an object-oriented, continuous-time, full simulation model for addressing a wide range of clinical, procedural, administrative, and financial decisions in health care at a high level of biological, clinical, and administrative detail. The full model has two main parts, which with some simplification can be designated “physiology models” and “models of care processes.” The models of care processes, although highly detailed, are mathematically straightforward. However, the mathematics that describes human biology, diseases, and the effects of interventions are more difficult. This paper describes the mathematical formulation and methods for deriving equations, for a variety of different sources of data. Although Archimedes was originally designed for health care applications, the formulation, and equations are general and can be applied to many natural systems. © 2002 Elsevier Science (USA). All rights reserved.

Keywords: Decision making health services administration models; Biological patient care model; Mathematical continuous time object-oriented simulation

1. Introduction

Simulation models are distinguished from other types of conceptual models by the fact that they include simulated objects, such as people, that correspond to real objects, one-to-one. Simulation models vary greatly in their breadth, depth, and realism. Our objective was to design a very broad, deep, and realistic model that could be used to address a wide range of clinical, administrative, and financial decisions in health care, at the level of detail at which real decisions are made. Development of such a model requires creating a population of simulated individuals who have all the important events that occur in real people and who respond to interventions in the same way as real people. In health care, this requires modeling all the essential aspects of human anatomy, physiology, pathology, and response to medical treatment. Because timing is an essential element of the oc-

currence, manifestation, progression, management, and outcomes of diseases, the model must also be continuous.

This paper introduces the mathematical formulation of such a model. The full model has two main parts, which with some simplification can be called the physiology models and the models of care processes. The models of care processes, while extremely detailed, are straightforward from a mathematical point of view. In contrast, the equations describing human physiology and diseases are much more difficult and are the subject of this paper. Specifically, the paper describes how we model human physiology and pathophysiology and how the necessary equations can be derived from different types of data sources or research studies. Although this model was originally developed for health care, its formulation and equations are general and can be applied to many natural systems.

2. Model formulation and definitions

The model is designed for object-oriented programming. The major classes of objects in the model include

* Corresponding author. Fax: +970-429-0882.

E-mail address: eddyaspen@yahoo.com (D.M. Eddy).

people (members/patients), facilities, personnel, interventions, equipment, supplies, records, policies, and budgets. This section describes how those models are formulated, and the main definitions.

To avoid confusion, we will call the individuals in the model “agents.” In the model, agents have physiologies and organ systems (e.g., heart, kidney, blood system, and immune system) just as real people do. (Agents also have many other attributes—e.g., names, locations, behaviors, and education levels—which are not of concern here.) In the model, as in reality, organ systems have “parts” and “functions.” Parts can have subparts and subparts can themselves have subparts. For example, in the model, the heart is composed of four coronary arteries, heart muscle, valves, and other parts; each coronary artery has multilayered walls and lumens (channels); the lumens can be occluded by a thrombus at any point.

All of these parts and subparts have functions that correspond to those present in real people. For example, the function of a coronary artery is to carry blood to the heart muscle and the function of the heart is to pump blood to other organs. As is true for real human organs, a successful functioning of agents in the model depends on all the agent’s parts and subparts functioning successfully.

The physiology of an agent is characterized by what we call “features,” which in health care correspond to a wide variety of anatomic and biologic variables. Examples of features in the model are blood pressure, cholesterol levels (i.e., high-density lipoprotein [HDL] and low-density lipoprotein [LDL]), bone mineral density, patency of a coronary artery, electrical potentials of the heart (as recorded on an electrocardiogram), contractility of myocardium, cardiac output, visual acuity, and serum potassium level. A feature can be continuously observable (e.g., a rash), intermittently observable through tests (e.g., diameter of a coronary artery), or not directly observable, except through resultant events (e.g., “spread” of a cancer).

At any time, every feature has a “value” (e.g., systolic blood pressure of 140 mm Hg, 75% occlusion in the left main coronary artery). Over time, the values of features change. This progression causes every feature in every individual to have a “trajectory,” which is defined mathematically as the value of the feature as a function of time.

As in reality, the trajectory of a feature in a particular agent can be affected by the agent’s characteristics, behaviors, and other features. In health care, these are often called “risk factors.” For example, the occlusion of a coronary artery can be affected by an individual’s family history (genetics), sex, age, use of tobacco, blood pressure, LDL cholesterol level, and many other risk factors.

If no interventions are applied to change it, the trajectory of a feature is called its “natural trajectory” or, in the medical vernacular, its “natural history.”

As in reality, when one or more features are considered “abnormal,” we say that an agent has a “disease.” Because in real life concepts of abnormality can change, definitions of diseases can change. Furthermore, many definitions of diseases are “manmade” and represent gross simplifications of the underlying physiology, and many diseases have different definitions put forth by different experts. Therefore, what people call a disease is actually a label we apply to a constellation of biological variables according to particular rules. For example, a person is said to have “diabetes” if their fasting plasma glucose exceeds 125 mg/dL or if their oral glucose tolerance test exceeds 199 mg/dL. For these reasons, we consider it important to model the underlying features rather than whatever definition of a disease is current. This approach is not only flexible, but also addresses comorbidities in a natural way through their underlying biological variables. When we talk about a “disease model,” such as the “diabetes model,” we are actually referring to the set of biological variables and interventions that are pertinent to that disease.

At any time, the values or progression of many features can be observed by “tests.” Examples are electrocardiograms, blood chemistry panels, and X-ray imaging. In this context, taking a patient’s history and conducting a physical examination are also tests. In health care, depending on whether a test is applied before the occurrence of any symptoms, after the occurrence of symptoms but before the diagnosis of the disease, or after the diagnosis of the disease, the use of a test is called “screening,” “diagnosis,” or “monitoring.”

As a feature progresses, it can cause certain “events” to occur. In health care, these events are typically called “signs” (which are measured by tests), “symptoms” (which are directly experienced or felt by the patient, such as a headache), or “health outcomes” (which are the major symptoms and consequences of a disease such as angina, heart attacks, and death). If an event is a direct manifestation of a biologic feature (e.g., “mild hypertension” or “hypercholesterolemia”), the event is called a “biologic event.” If an event is not a direct manifestation of a feature but is caused by or associated with a feature, as in angina caused by partial occlusion of a coronary artery, we will call it a “clinical event.”

For many diseases, there are “health interventions” that can change the value of one or more features, the rate of progression of one or more features, or both the value and rate of progression. Interventions can affect features either indirectly (by changing risk factors, e.g., smoking) or directly (by changing the feature itself). Health interventions that have direct effects can change either the value of a feature (e.g., performing bypass surgery to open an occluded coronary artery) or the rate

of change of a feature (e.g., lowering cholesterol level to slow the rate of occlusion). In addition to health interventions, there are “logistic interventions” that change the way care is delivered. “Care processes” (e.g., increasing the use of case managers, altering referral criteria for specialists, or extending a clinic’s hours of operation) can also change the way care is delivered. In this paper, we will be concerned only with health interventions and will use the term “intervention” in that sense.

The occurrence of signs, symptoms, and outcomes can set in motion a wide variety of “logistic events” in the health care system. These in turn involve other classes of objects in the model. For example, when chest pain causes a person to call a hospital (a “behavior”), the person who answers the phone (a type of “personnel”) determines the seriousness of pain (by applying a particular triage “policy”). The patient may go to an emergency department (a type of “facility”) where he or she is seen first by an ED nurse (another type of personnel), is asked to relate a brief medical history, and receives a physical examination and then electrocardiography (ECG) (all of which are “tests”). Moreover, electrocardiography uses an ECG machine (an item of “equipment”) and ECG paper (a type of “supply”).

The core of the simulation model is the set of equations that describe the physiologies of individuals. These include the equations that describe: (a) natural trajectories of all the important features, including interactions between features and the effects of other risk factors, (b) the occurrence of clinical events as a function of features, (c) the effects of interventions on features and on clinical events, and (d) the functions of organs. This paper describes a general method for deriving equations for the first two of these, based on existing data. The functions of organs and the effects of interventions are best described in terms of specific diseases and are only summarized here.

3. Equations for trajectories of features

The equations for the features that define important diseases depend on the number of features, the number of events, and the available data. We will begin with the simplest case in which a person has a single feature and there are person-specific data on the values of the feature at a series of times. For example, think of the organ being the heart, the part of the organ being a coronary artery, the feature being the degree of occlusion of the artery, and an event associated with the feature being a heart attack. Later, we will discuss strategies when the data are less good and extend the model to include additional features, dependence of features on risk factors (including other features), and the effects of interventions.

For each agent, we want to define a function that describes the natural progression or trajectory of the feature over time, from birth to death, where “natural” means the trajectory of the feature in the absence of any special interventions from the health care system. Other equations can then be used to simulate the effects of interventions.

Index a particular agent by k and let the trajectory of the feature for the k th agent be $F^k(t)$, where t is the time since the agent’s birth (age). Because interventions can change either the value of a feature or the rate of change of a feature, we want to write a differential equation for $F^k(t)$. The general form of the differential equation for each agent is

$$\frac{dF^k(t)}{dt} = R^k(t), \quad (1)$$

where $F^k(t)$ is the value of the feature at time t for the k th agent and $R^k(t)$ is the rate at which the value of the feature is changing at time t (the derivative). Either $F^k(t)$ or $R^k(t)$ determines the natural trajectory for the k th agent and either $F^k(t)$ or $R^k(t)$ can be determined from the other. To simplify the description of the methods, we focus here on the value of the feature, $F^k(t)$, with the understanding that the rate of change of feature, $R^k(t)$, can always be derived from $F^k(t)$ by Eq. (1).

4. Trajectories of features as a random process

Our goal is to create a set of trajectories for a population of simulated agents that statistically match the trajectories of a population of real people. We formulate the problem as follows. Consider the trajectories of real people to be a random (stochastic) process parameterized by age (t). (As discussed later, the random process can also be made conditional on risk factors, including other features.) The sample space for the random process for a particular feature is the collection of all possible trajectories, one for each person. Call this space Ω , with elements $\omega = \{\omega_1, \omega_2, \omega_3, \dots\}$, where ω_k specifies the trajectory of the feature for a particular person. In this space, the random process for the trajectories is designated by upper-case boldface font, $\mathbf{F}(\omega, t)$. Each function in Eq. (1) is a realization of the stochastic process. That is, $F^k(t) = \mathbf{F}(\omega_k, t) = \omega_k$ all represent the trajectory of the k th person in the set ω .

The next step is to derive a distribution for the random process from observations of real people. Once this distribution has been derived, the third step will be to draw trajectories from the distribution at random to create new, simulated, agents. This process will guarantee that the trajectories of the simulated population will statistically match the trajectories of the real population.

The process of estimating a distribution for the random process involves two main steps: First, use person-specific data to derive the samples that will define the distribution and then determine the distribution from the samples.

5. Deriving samples for estimating the distribution for the random process

The general method for estimating a distribution for the random process begins by writing the equation for the trajectory of the feature as an expansion, each term of which consists of a basis function weighted by a coefficient. For each real person, we will find the specific values of the coefficients that provide the best fit to that person's data. Those values will become the samples from which to derive the distributions for the coefficients, which in turn will determine the distribution for the random process.

There are many ways to estimate the specific values for the coefficients, depending on how the basis functions are chosen. Each method has strengths and weaknesses. We will begin by describing a method based on Fourier expansions that uses standard mathematical techniques and is guaranteed to converge. To illustrate the range of possible methods, we will also describe a more intuitive method that is more closely related to the familiar regression techniques used to analyze health data. Unlike the Fourier expansion, the latter method, which we will call the hybrid expansion, is not guaranteed to converge.

6. Determining samples by using the Fourier expansion

To derive the samples of the distribution for the random process from person-specific data using a Fourier expansion, expand the random process $F(\omega, t)$ (or any function of $F(\omega, t)$ such as the log of the odds ratio of $F(\omega, t)$) in a Fourier-type series. Each term of the series consists of two parts: an age-dependent, deterministic (nonrandom) "basis" function (denoted as $P_j(t)$ for the j th term in the expansion), multiplied by a coefficient that is a random variable (denoted by a lower-case boldface letter) and is independent of age, $f_j(\omega)$. Thus,

$$F(\omega, t) = \sum_{j=0}^{\infty} f_j(\omega) P_j(t). \quad (2)$$

The basis functions $P_j(t)$ could be any complete set of functions. Some examples are: a polynomial series, i.e., t^j ; the j th Legendre or Laguerre polynomial; or a classical Fourier series, i.e., $\sin(jt/T)$. When the basis functions are chosen to be orthonormal over the range of ages of interest, then the expansion is called a Karhunen–Loeve (K–L) decomposition [1–5]. Because

the theory of K–L decompositions is reasonably well developed and because the K–L decomposition has several advantages, there are good reasons to choose the $P_j(t)$ to be orthonormal. The Legendre, Laguerre, and Fourier functions are all orthonormal.

Whatever basis functions are chosen, they will be the same for every person or agent. However, because the coefficients $f_j(\omega)$ are random variables, they are different for each person or agent. The choice of basis functions thus affects the coefficients calculated and the rate of convergence for the series (i.e., number of terms needed to fit the data) but will not prevent the method from working.

The task now is to estimate the samples that will be used to derive the distributions for the coefficients $f_j(\omega)$. In practice, the sum in Eq. (2) is truncated to a finite number of terms, $J + 1$, which is related to (but not greater than) the number of events observed for each person. The method for estimating the samples for the $f_j(\omega)$ depends on the available data. In the best case, there are person-specific data that provide a series of values of the feature at specified times for a large number of people. For example, there might be a series of measurements of intraocular pressures for a group of people. There is no requirement that the measurements for each person be taken at the same times.

The function describing the trajectory for the k th real person is approximated by a finite sum,

$$F^k(t) \approx \sum_{j=0}^J f_j^k P_j(t), \quad (3)$$

where f_j^k are the coefficients determined to fit the data observed for the k th person. The f_j^k are the samples that will be used to estimate the distribution of the coefficients $f_j(\omega)$. Although many different ways can be used to estimate the f_j^k from the data, we outline three methods here: (a) the first requires the expansion in Eq. (3) to pass through all of the observed points, (b) the second uses the method of least squares, and (c) the third uses the orthonormal properties of $P_j(t)$.

For the first method, imagine that for each person there are $J + 1$ observations. This will lead to $J + 1$ equations with $J + 1$ unknowns, the f_j^k . This linear system of equations can be solved for the f_j^k using standard methods.

The second method, least squares, is useful when the number of terms is less than the number of observations for each person. For example, if there are M observations that can be used to determine coefficients for the $J + 1$ terms, where $J < M$, then the f_j^k can be determined by minimizing the sum of the squares of the differences between the value of the function and the value of the expansion on the right-hand side of Eq. (3) at all of the M points. The expression to be minimized for this method, for the k th person, is

$$\sum_{m=1}^{m=M} \left(F^k(t_m) - \sum_{j=0}^{j=J} f_j^k P_j(t_m) \right)^2 \quad (4)$$

Taking the derivative of this equation with respect to each f_j^k ($j = 0$ to J) and setting this derivative to zero produces a set of linear equations that determine the f_j^k .

The third way to determine the f_j^k makes use of the orthonormal properties of the $P_j(t)$. Multiplying both sides of Eq. (3) by $P_i(t) * W(t)$ (where $W(t)$ is the weight for the chosen set of orthonormal functions) and using the orthogonality property directly yield the following expression for f_j^k :

$$f_j^k = \int F^k(t) * P_j(t) * W(t) dt. \quad (5)$$

The observed points are used to approximate the integral. As before, there must be at least $J + 1$ observations. The coefficients determined in this way will minimize the integral of the square of the difference between the right and left sides of Eq. (3). That is, the coefficients calculated by Eq. (5) will minimize

$$\int \left(F^k(t) - \sum_{j=0}^{j=J} f_j^k P_j(t) \right)^2 W(t) dt. \quad (6)$$

The theory of this type of expansion is called functional analysis. An important advantage of this method is that the power of the theory of functional analysis can be applied to the estimation procedure. Moreover, many properties of the K–L decomposition require the use of this type of expansion.

For any set of basis functions chosen initially, any of these three methods can be used to find values of the coefficients that cause each person’s trajectory to fit the data. At this point, we can create “clones” of the original population, i.e., agents whose trajectories match, person for person, the trajectories of real persons. These agents could be used to explore outcomes and effects of interventions in the original population.

As valuable as that might be, however, creating only clones would not enable us to simulate other populations different in size, risk factors, and other characteristics. To do that, we need to create agents that statistically match, but are not identical to, the real people. This is accomplished in three steps: (1) for each coefficient, calculate the values of f_j^k for a large number of real people; (2) use these values to estimate distributions for each $f_j(\omega)$; (3) draw new values from the distributions to create new simulated individuals (agents).

7. Ensuring first-order independence of $f_j(\omega)$

Before proceeding to calculate distributions for the $f_j(\omega)$, however, it is important to determine if they are independent, at least to the first order. If they are, a

particular agent could be created by drawing values for each of the j random variables $f_j(\omega)$ and then using Eq. (3) to calculate a particular simulated trajectory. But it is very unlikely that the $f_j(\omega)$ are independent. The presence and degree of covariance or correlation between the $f_j(\omega)$ can be determined and appropriate corrections can be made, as follows.

If the $f_j(\omega)$ are independent, then their covariance should be zero. To determine if this were true, we first transform the values of f_j^k for each person by subtracting out the mean of the values of the coefficient. (We represent the mean of a coefficient with angle brackets.) Thus, for the j th coefficient

$$\langle f_j \rangle = \frac{1}{K} \sum_{k=1}^K f_j^k, \quad (7)$$

where K is the total number of individuals for which data exist. Then for the k th individual, subtracting out the means of each coefficient in Eq. (3) yields

$$F^k(t) = \left(\sum_{j=0}^J (f_j^k - \langle f_j \rangle) P_j(t) \right) + \left(\sum_{j=0}^J \langle f_j \rangle P_j(t) \right). \quad (8)$$

The coefficient of the first term on the right is the original coefficient with the mean subtracted out. The last term on the right is required to maintain the equality of the two sides, and can be thought of as the average trajectory—the basis functions weighted by the average values of the coefficients. We will represent the average trajectory as $\langle F(t) \rangle$. That is,

$$\langle F(t) \rangle = \sum_{j=0}^J \langle f_j \rangle P_j(t). \quad (9)$$

Let q represent the transformed values for the coefficient; that is, for the j th coefficient and the k th person

$$q_j^k = f_j^k - \langle f_j \rangle. \quad (10)$$

This gives a new equation for the trajectory of the feature. Substituting Eqs. (9) and (10) in Eq. (8) yields

$$F^k(t) = \sum_{j=0}^J q_j^k P_j(t) + \langle F(t) \rangle. \quad (11)$$

Now define the covariance matrix C with elements C_{ij} , where

$$C_{ij} = \frac{1}{K} \sum_{k=1}^K q_i^k q_j^k. \quad (12)$$

If the random variables for the original coefficients $f_j(\omega)$ are independent, then the off-diagonal terms of the covariance matrix will be 0. In such cases, it is appropriate to create new simulated agents by drawing values from the distributions of the random variables for the coefficients and using these values in Eq. (3) to derive simulated trajectories for as many agents as desired.

When the original coefficients are *not* independent, which is the usual case, two main approaches are possible: (a) estimate a joint distribution for the $f_j(\omega)$ and create simulated agents by drawing from that joint distribution; or (b) use the covariance matrix to determine a new set of basis functions, $Q_j(t)$, and new coefficients, $s_j(\omega)$, that are not correlated. (The covariance is zero.) The latter approach requires fewer data, is computationally simpler, creates an optimal expansion, and can provide powerful insights into the behavior of the feature. This approach is closely related to both the principal component method (PCM) and the method of factor analysis and is a central feature of the K–L decomposition. After the new, uncorrelated coefficients $s_j(\omega)$ have been determined, if there are any higher-order correlations, then it is much easier to estimate their joint distribution and draw from that distribution to create simulated agents. (Under some conditions, the new coefficients will also be independent.)

The second approach is accomplished as follows. Because the covariance matrix is real, symmetric, and nonnegative, it has $J + 1$ real eigenvalues λ_j (with $\lambda_j \geq 0$) and $J + 1$ orthonormal eigenvectors ψ^j . The eigenvectors and eigenvalues have two important properties. First, multiplying an eigenvector by the matrix from which it was derived reproduces the eigenvector scaled by the eigenvalue. Thus,

$$\sum_{l=0}^J C_{jl} \psi_l^n = \lambda_n \psi_j^n \quad (j = 0 \dots J, \quad n = 0 \dots J). \quad (13)$$

Second, the eigenvectors are orthonormal,

$$\sum_{j=0}^J \psi_j^n \psi_j^l = \delta_{nl}, \quad (14)$$

where $\delta_{nl} = 0$ if $n \neq l$, and $\delta_{nl} = 1$ if $n = l$. Moreover, the eigenvectors span the space so that any vector can be represented as the sum of coefficients times the eigenvectors.

Using the eigenvectors of the covariance matrix, it is possible to calculate new coefficients and basis vectors for the expansion of the trajectory that have the desired property that the coefficients are uncorrelated. The first step in this calculation is to expand the coefficients q_j^k calculated for each person from Eq. (10) in terms of the eigenvectors and new coefficients s_i^k ,

$$q_j^k = \sum_{i=0}^J s_i^k \psi_j^i. \quad (15)$$

Eq. (15) can then be used to solve for the s_i^k in terms of the q_j^k . Multiplying each side by the n th eigenvector and summing over its elements yield

$$\sum_{j=0}^J q_j^k \psi_j^n = \sum_{j=0}^J \sum_{i=0}^J s_i^k \psi_j^i \psi_j^n. \quad (16)$$

But by Eq. (14) and the orthogonality of the eigenvectors,

$$\sum_{j=0}^J \sum_{i=0}^J s_i^k \psi_j^i \psi_j^n = s_n^k. \quad (17)$$

This equation defines the new coefficients in terms of the q_j^k and the eigenvectors; the new coefficients are a linear combination of the old coefficients and are weighted by the elements of the corresponding eigenvectors. Thus, for the n th new coefficient, we obtain

$$s_n^k = \sum_{j=0}^J q_j^k \psi_j^n. \quad (18)$$

Similarly, we can define new basis vectors $Q_j(t)$ as linear combinations of the old basis vectors weighted by the elements of the eigenvectors. That is,

$$Q_n(t) = \sum_{j=0}^J \psi_j^n P_j(t). \quad (19)$$

Using Eq. (18) we can verify that the coefficients $s_j(\omega)$ and $s_n(\omega)$ are not correlated. Thus,

$$\begin{aligned} \langle s_j(\omega) s_n(\omega) \rangle &= (1/K) \sum_{k=1}^K \left(\sum_{i=0}^J q_i^k \psi_i^j \right) \left(\sum_{l=0}^J q_l^k \psi_l^n \right), \quad (20) \\ &= \sum_{i=0}^J \sum_{l=0}^J C_{il} \psi_i^j \psi_l^n = \sum_{i=0}^J \lambda_n \psi_i^j \psi_i^n = \lambda_n \delta_{jn}. \end{aligned} \quad (21)$$

Further, by substituting the new coefficients and basis functions, we can verify that these new coefficients and basis functions satisfy the original equation for the trajectory of the feature. Substituting Eq. (15) in Eq. (11) thus yields

$$F^k(t) = \langle \mathbf{F}(t) \rangle + \sum_{j=0}^J \sum_{l=0}^J s_l^k \psi_j^l P_j(t) \quad (22)$$

and substituting Eq. (19) in Eq. (22) yields

$$F^k(t) = \langle \mathbf{F}(t) \rangle + \sum_{l=0}^J s_l^k Q_l(t). \quad (23)$$

To summarize, starting from an arbitrary set of basis functions, $P_j(t)$, this method can be used to derive a set of basis functions, $Q_j(t)$, that cause the trajectories of real persons to best fit the observed data (e.g., passing through all observed points), but for which the coefficients, $s_j(\omega)$, are uncorrelated.

This method of expansion has many advantages. First, it corrects for first-order correlations. If the random process is Gaussian, then correcting for first-order correlations corrects for all higher-order correlations and, consequently, makes the random variables $s_j(\omega)$ independent. (Loosely speaking, the random process $\mathbf{F}(\omega, t)$ is Gaussian if for an arbitrary set of times t' the

$F(\omega, t')$ are jointly Gaussian.) Although assuming a Gaussian distribution is frequently reasonable, the method does not correct for higher-order correlations. If higher-order correlations are found to be important, then forming the joint distribution of the $s_j(\omega)$ may still be necessary. Even in this case, however, forming these joint distributions from Eq. (23) will still be easier because the first-order correlations would have been removed.

A second advantage of this method is that it provides insight into the nature of the trajectory of the feature. The K–L expansion can be shown to be optimal in the sense that if the expansion in Eq. (2) is truncated at the m th term, the mean square error is smallest if the basis functions are the $Q_j(t)$ and the coefficients of the expansion are the s_j^k as derived above. By exploring the rate at which the expansion converges when different basis functions are used and by exploring the behavior of the components of the expansion's trajectory (the $Q_j(t)$) we can learn about the biology of the feature. Also, using the new basis functions is likely to make the expansion converge faster in the sense that fewer terms are needed to get a good fit of the data. This can provide information about the minimum number of observations needed to formulate an accurate description of the feature's trajectory: the number of data points needed is equivalent to the number of expansion terms that have important coefficients. For example, if the data are well fitted by using only two terms in the expansion, only two data points will be needed to fit the entire function. This fact can have important consequences for future data collection.

We can assess the importance of each term in the expansion by examining the sizes of the eigenvalues λ_n . This process is similar to factor analysis. The covariance matrix has diagonal elements σ_n^2 , where $\sigma_n^2 = 1/K \sum_{k=1}^K (q_n^k)^2$. The sum of the diagonal elements of \mathbf{C} is $\sigma^2 = \sum_{n=1}^N \sigma_n^2$. This sum is conserved in diagonalization, so the sum of the eigenvalues is also σ^2 . Just as in factor analysis, the size of each eigenvalue represents the importance of each term in the expansion of the process; the terms with the largest eigenvalues contribute the most to the convergence of the series. Consequently, the number of terms in the expansion can be reduced by keeping only those that have the largest eigenvalues. One frequently used method is to order the eigenvalues by size, calculate their sum, and retain the first m eigenvalues such that $\sum_{i=0}^{i=m} \lambda_i \geq \text{Frac} * \sigma^2$, where Frac is the percentage of the original variance the reduced eigenvector set will reproduce. Frac is typically chosen to be about 0.9. Standard (but nonetheless empirical) methods of choosing the number of eigenvalues to retain in the factor analysis method are described elsewhere [6,7].

In summary, the Fourier expansion with the K–L decomposition produces a new set of coefficients that are

easier to use because they are uncorrelated (and perhaps independent). If higher-order correlations exist, then the K–L procedure makes finding the joint distribution of the coefficients easier. In addition, because the expansion is optimal, fewer terms in the series may be needed to adequately represent the random process. The K–L procedure also enables identification of terms to be retained. A drawback of the Fourier expansion is that the terms of the expansion equations, either before or after transformation, are abstract and do not correspond to real biological variables. They will calculate the trajectory of a feature accurately, but will have little if any intuitive or biological interpretation.

8. Determining samples using the hybrid expansion

In some cases, the use of some nonstandard functions may be helpful as part of the set of basis functions. For instance, when it is reasonable to believe that a feature depends strongly on one or more other features, it would be natural to try to incorporate that dependence explicitly into the basis functions. For example, occlusion of a coronary artery (F_1) is known to depend on both blood pressure (F_2) and cholesterol level (F_3), among other features. These features can be included in the expansion for F_1 as follows:

- (a) As before, denote the set of basis functions as $P_j(t)$. However, instead of choosing the $P_j(t)$ to be orthonormal, let $P_0(t)$ represent the blood pressure level and let $P_1(t)$ represent the total cholesterol level for that person. Additional basis functions could be chosen to address dependencies or other relations between features. For example, we could let $P_2(t)$ represent the product of blood pressure and total cholesterol levels and let $P_3(t)$ represent the product of three values: t , blood pressure level, and cholesterol level. As in the Fourier expansion, the remaining basis functions would be the orthonormal set.
- (b) After the first few basis functions are chosen to include other features, the remainder of the analysis can proceed as for the Fourier expansion, except that Eq. (5) cannot be used to determine the coefficients (i.e., because the full set of basis functions is no longer orthonormal). However, the other equations will still apply. For example, the covariance matrix can still be diagonalized to obtain a new set of basis functions having the desired properties. Notice, however, that the first few basis functions will be different for every person because the functions describe the progression of a particular feature for a particular person.

This type of hybrid expansion is related to the expansions traditionally used in regression analyses. The independent variables in a regression equation correspond to the basis functions in our model and the

coefficients correspond to our coefficients. However, there are important differences: the basis functions in our model are usually functions of time and our coefficients are random variables. Nonetheless, both models explicitly include other features as terms in the expansion.

The hybrid expansion has several advantages: (a) it is intuitively appealing; (b) unlike the Fourier expansion, its basis functions correspond to biological variables and have immediate clinical interpretations; (c) it corresponds to regression models, which are familiar; and (d) it can determine how important is the dependence of one feature on another (e.g., the importance of blood pressure level in determining the progression of coronary artery occlusion). Moreover, the hybrid expansion can converge even faster than can the Fourier expansion.

However, the hybrid expansion also has several disadvantages: (a) whereas the Fourier expansion is guaranteed to converge under very general conditions, the hybrid expansion is not; (b) the hybrid expansion is not as mathematically rigorous as the standard method; and (c) some matrices of the hybrid expansion may be ill conditioned and difficult to invert. Nonetheless, the hybrid expansion may be a useful practical approach, especially when data or existing regression models describe the dependence of one feature on other features. Because data were available from the Framingham study [8], we chose the hybrid expansion for modeling coronary artery occlusion.

9. Using samples to determine the distribution for the random process

Determining the distribution of a random variable from a set of samples (s_{ij}^k) is a standard problem that is often addressed using maximum likelihood techniques. We first review the application of this technique for a feature that does not depend on any other features. We then show how to include dependence on other features.

Designate the samples as s_{ij}^k , where k represents the k th individual, j represents the j th term in the expansion, and i represents the i th feature. We denote the probability distribution of the random variable for the samples, $s_{ij}(\omega)$, as ρ_{ij} . ρ_{ij} is characterized by a small number of parameters:

$$\begin{aligned} \rho_{ij}(x|\theta_1^j, \theta_2^j, \dots, \theta_N^j) dx &= \rho_{ij}(x|\vec{\theta}^{ij}) dx \\ &= P(x < s_{ij}(\omega) < x + dx). \end{aligned} \quad (24)$$

In this equation, $P(\dots)$ is the probability that the random variable $s_{ij}(\omega)$ lies in the range between x and $x + dx$, conditional on the $\vec{\theta}^{ij}$, where $\vec{\theta}^{ij} = \{\theta_n^j, n = 1 \dots N\}$ are the parameters of the distributions of the coefficients $s_{ij}(\omega)$. The probability of obtaining the samples s_{ij}^k is the

likelihood and is related to the distribution ρ_{ij} and to the samples s_{ij}^k by the likelihood function

$$L(\vec{\theta}^{ij} | s_{ij}^1, s_{ij}^2, \dots, s_{ij}^K) = \prod_{k=1}^K \rho_{ij}(s_{ij}^k | \vec{\theta}^{ij}). \quad (25)$$

An estimate of the parameters $\vec{\theta}^{ij}$ is obtained by maximizing the likelihood as a function of the parameters $\vec{\theta}^{ij}$.

10. Incomplete data

The methods just described can be applied when there are person-specific data for the values of the feature for several times (but not necessarily at the same times for each person.) This is realistic for many problems today and is a restriction shared by most statistical models, such as regression models. Moreover, person-specific data are likely to become far more available with increased use of automated clinical information systems.

Today, however, there are many clinical conditions for which a biological feature is difficult or impossible to observe and for which the only data available relate to occurrences of clinical events. For example, several large epidemiologic studies provide data on the probability of a heart attack for people of various ages, but no large studies provide data on the degree of occlusion of coronary arteries (because the required measurement entails use of a risky, expensive test). In such cases, the choice of an approach depends on the availability of data from ancillary sources on the relationship between the feature and a clinical event. For example, Roberts [9] has reported the degree of occlusion in patients who recently had a heart attack. When available, data like these can be used to translate epidemiologic data on clinical events into estimates of values of the feature, and the process described can be used to complete the derivations of equations for the trajectory of the feature.

When there are no data at all on the value of the feature at the time of clinical events, a different approach can be used. In this case, the objective is no longer to write equations for the trajectory of the true values of the feature, because that is not possible if there are truly no systematic observations of the feature. Instead, the objective can be thought of as to write equations for an *imaginary* or *virtual* feature whose only purpose is to accurately reproduce the observed occurrences of clinical events. For this purpose, the feature can be assigned an arbitrary value when the event occurs. If there are more than one clinical events to be simulated, the arbitrary values should obviously correspond to the order in which the events occur. If the events occur in different orders in different people, chances are good that the events are caused by different features, and equations for each feature can be derived

accordingly. Although this approach is worthless for providing information about the true value of the virtual feature, it does provide what is needed for an accurate simulation of clinical events at rates that match the occurrences of real clinical events.

The last case arises when there are no person-specific data and the only available data are aggregated over a population. For example, there may be data on the age distribution of patients diagnosed with various stages of a cancer, but no person-specific data on the ages at which particular individuals pass through each stage. Of course, if there are data from other sources that relate the clinical events to the values of the feature (in this example the stage of the cancer), those data should be used to resolve the problem as described in the previous section. Assuming that there are no such data, there are two main options, depending on whether there is reason to believe that the clinical events are correlated. If it is reasonable to assume that they are not correlated, then they can be modeled as if caused by two different features, and the modeling problem is reduced to a case we have already discussed. If it is not reasonable to assume that the events are uncorrelated, then we have to postulate a model that describes the correlation. We would first look for any data on which our presumption of correlation was based and use those data to develop a model. But even if no such data are available, there may be plausible reasons to postulate a model. For example, we might assume that some individuals have an “aggressive” form of cancer, implying that they will move through each stage relatively rapidly, whereas others may have a more “indolent” form, implying that their cancer will tend to progress more slowly. Thus, if a person with aggressive disease was in the first 10% in terms of the age at which the first stage of the disease developed, we might be willing to assume that he will be in the first 10% in the pace at which he progresses through subsequent stages. Once a specific correlation is postulated, then it is possible to convert the cross-sectional data into a set of person-specific longitudinal data and the problem is transformed into the original case.

It is important to emphasize that whether or not assumptions of the type described in this section must be made, the ultimate test of the quality of the model will be whether it can reproduce whatever data exist. If it can, the model should be useful. If it cannot, the model should not be used for a problem involving that feature until better data are available. These issues are addressed in the validation of the model.

11. Multiple features and interdependence of features

The equations for multiple features depend on the extent to which features are independent in the sense

that they depend only on time (e.g., a person’s age) and do not depend on other features or other factors that may vary across individual persons. It should be apparent that for features that are independent in this sense and depend only on an individual’s age, the methods already described can be used to derive equations for as many such features as desired. The difficulties arise when the trajectory of a feature depends on other features or other risk factors.

This is a common occurrence. For the example of coronary artery disease, the rate of coronary artery occlusion depends not only on age but also on other features, such as cholesterol level, blood pressure level, tobacco use, and diabetes. Collectively, we call these “risk factors,” with the understanding that this term covers a wide range of disparate factors. Some such factors are fixed characteristics (e.g., sex, race), some are biologic features (e.g., cholesterol level), some are behaviors (e.g., smoking), some can be modified by interventions, and some cannot. Fortunately, the method for incorporating risk factors in the trajectory of a feature works for all types of risk factors. In the following discussion, we will refer to incorporating a dependence on features, with the understanding that the method can easily incorporate dependence on other risk factors.

Begin by observing that the dependence of one feature on other features is already incorporated in the data and therefore is incorporated in the coefficients and basis functions estimated for each individual in Eqs. (3), (11), or (23). Our task is to separate out that dependence and to represent it explicitly in the coefficients or basis functions of the equations for the trajectory of the feature. We need to do this if we are to develop a general model that can be used to analyze interventions, not only in clones of the original population, but also in a wide variety of other populations that will have different distributions of risk factors.

The separation of the dependence on other features requires care, because the data for estimating the equations for a feature contain all the dependence of the feature on age. But the data are not separated into the dependence of the feature as a function of age, at a fixed value of another feature, or the dependence of the feature as a function of another feature, at a fixed age.

The dependence we want to address can be represented either in the coefficients or in the basis functions. In the hybrid expansion, the dependence is represented in the basis functions or in both the basis functions and the coefficients. In the hybrid expansion, it is easy to include the dependence of one feature on another because the independent features (such as cholesterol level in the trajectory for coronary artery occlusion) can be explicitly separated out and included in the expansion as its own basis function. The trade-off is that the hybrid expansion is not guaranteed to converge and the equa-

tions for determining the coefficients for the hybrid expansion may be ill conditioned. In the Fourier expansion, the dependence is represented in the coefficients. Here, we will describe how to determine the distributions of the coefficients from the available data, when the features are related in a Fourier expansion and one feature depends on another.

Drawing on the notation of Eqs. (24) and (25), we can consider the distributions of the coefficients of the random process for the i th feature, $\mathbf{F}_i(\omega, t)$, to depend on the coefficients of the random processes of other features. Let $\hat{s}_i(\omega)$ represent the coefficients of the expansion terms for the random processes of all features other than feature i (i.e., all $s_{i'j}(\omega)$ for $i' \neq i$ and all j), and let \hat{x}_i represent specific values that the coefficients can take for all features other than feature i . To allow the distributions of the coefficients of the i th feature to be conditional on the coefficients of other features, we represent the parameters of the distributions for the i th feature to be functions of the coefficients of the distributions of other features, that is, $\vec{\Theta}^{ij}(\hat{x}_i)$. Using this notation,

$$P(x < s_{ij}(\omega) < x + dx | \hat{s}_i(\omega) = \hat{x}_i) = \rho_{ij}(x | \vec{\Theta}^{ij}(\hat{x}_i)). \quad (26)$$

The $\vec{\Theta}^{ij}(\hat{x}_i)$ can be made functions of the coefficients \hat{x}_i in many different ways. One choice is to use an expansion that is linear in the coefficients,

$$\vec{\Theta}^{ij}(\hat{x}_i) = \vec{\Theta}^{ij} \left(\vec{\beta}_0^{ij} + \sum_{i' \neq i, \text{all } j'} \vec{\beta}_{i'j}^{ij} x_{i'j} \right). \quad (27)$$

In general, let \vec{B} be the coefficients that define the functional representation of $\vec{\Theta}^{ij}(\hat{x}_i)$. (If the functional representation is a linear expansion, then $\vec{B} = \{\vec{\beta}_0^{ij}, \vec{\beta}_{i'j}^{ij}\}$). The likelihood of obtaining all the sample values s_{ij}^k for all the individuals $k = 1 \dots K$, and all the features i , and the coefficients of all the terms of the expansion (indexed by j) is given by the equation

$$L(\vec{B} | \vec{s}) = \prod_{k=1, i, \text{all } j}^{K, I} \rho_{ij}(s_{ij}^k | \vec{\Theta}^{ij}(\hat{x}_i)), \quad (28)$$

where \vec{s} represents the set of all coefficients obtained by observations on all people. The \vec{B} are determined by maximizing the likelihood in Eq. (28).

In practice, most features are either independent or depend on only a few other features. In the latter situation, variables that depend on each other can be grouped together, isolated as a group, and estimated separately from the rest. The particular model chosen for the dependencies, whether it involves the Fourier expansion or the hybrid expansion, will be different for different features and health conditions, depending on the biology and the available data. How well the model addresses any dependencies between biological variables will ultimately be tested in the validation exercises.

12. Equations that relate events to features

After functions have been derived for the natural trajectories of features, linking features to events is fairly straightforward. First, because any biological variable can be represented in the model as a feature, biologic events are represented by the values of features; tests can be applied to measure the value of a biological feature at any time. Uncertainty, random errors, and systematic errors in tests are easy to include.

For clinical events, more work is required. If the clinical event is directly related to the feature (as, for example, occlusion of a coronary artery can be observed through an infarction), the trajectory will automatically reproduce the occurrence of the clinical event as required. Otherwise, it is necessary to describe or model how the clinical event is linked to the feature. The appropriate model will depend on the data available. For example, a standard medical text [10] suggests that angina pain tends to occur when degree of coronary artery occlusion approaches 70%. Clinical events can also be defined as more complex functions of one or more features. For example, rapid weight change in a patient with congestive heart failure is an indication to regulate the dose of diuretics. Because values of all features are always available in continuous form through the equations for their trajectories, it is relatively easy to define models that determine the occurrence of clinical events as functions of the values of several features. For example, clinical diabetes can be defined as either a fasting plasma glucose level >126 mg/dL or a glucose tolerance test result >200 mg/dL.

13. Effects of interventions

Effects of health interventions can be modeled either as a change in the value of a feature, the rate of change of a feature, or a combination of the two. The choice and the exact model will depend on the biology of the disease, the mechanism of action of the intervention, and the available data.

14. Application of the methods

The methods described in the previous sections can be used to estimate the trajectories of features, for a variety of different types of available evidence. In practice, a model for a specific disease is developed as follows.

1. We begin with a description of the physiology and pathophysiology pertinent to the disease. This description is obtained from textbooks, review papers, and experts.
2. We next develop a nonquantitative, conceptual model of the disease that identifies all the important

biological variables and their relationships. This can be represented by a figure with circles representing the variables and lines or arrows representing the relationships between variables. At this time, we also identify the primary feature or features that “cause” the disease and that are responsible for its progression. One or more of the primary features may be virtual, depending on how well the causes of the disease are understood.

3. We next identify the physiological, clinical, and epidemiological research studies that form the basis for the current knowledge of the disease.
 - a. Ideally, there will be a database that contains longitudinal, person-specific data on all the important biological variables and clinical outcomes. When such data exist, either the Fourier expansion or the hybrid method can be used to derive equations for the features. The resulting equations will incorporate any dependencies between features that can be identified in the data.
 - b. When person-specific data are not available for a disease, we use the studies that form the current knowledge base for the disease. Typically, there are dozens of studies of various designs that address particular variables or sets of variables and their relationships. For example, a large epidemiological study might provide information on the age- and sex-specific incidence rates for various racial and ethnic groups; a regression equation from another large database might give the effects of biological risk factors; a clinical trial might provide information about the effects of changing one of those biological variables; and so forth. Just as experts do cognitively, we will use these pieces to build a model of the disease by integrating the information from all the sources that experts consider to be important. The main differences are that our model will be quantitative and can be subjected to validation.
 - c. Even when person-specific data are available, there are invariably important aspects of the disease that are not completely addressed by the database. We use additional studies to complete the picture of the disease as just described.
4. The model is then ready for validation.

Throughout this process, we are guided by experts. They provide an initial description of the disease, determine the appropriate level of biological and clinical detail, help identify the research studies that form the knowledge base for the disease, help identify clinical trials and other studies that can be used for validations, help interpret the validations, and help identify new technologies and research results that should be incorporated in new versions of the model. However, we do not use any expert subjective judgments for variables in the model. All of the equations and parameters are

derived from at least one actual research study (except for “what-if” calculations).

15. Validation

Ultimately, the validity of everything—the basic formulation, the choice of objects to represent the anatomy and physiology, the use of virtual features, the methods for deriving equations, the modeling of dependencies between features, and the combination of results of different studies—is determined by how well the model can reproduce and predict empirical observations. We evaluate the validity of the model through four different types of exercises. The first is that each of the equations must fit the data used to derive them. These exercises help ensure that each individual piece of the model is an accurate representation of the results of a particular study, but they do not test the connections between pieces or the consistency of results across studies. We call these “one-star” validation exercises.

The other three types of validation exercises are far more challenging, and draw on the ability of an Archimedes model to simulate clinical scenarios at a high level of detail. They involve simulating clinical studies and comparing the simulated results with the real results. The basic steps for this type of validation are: (1) Have the model “give birth” to a large population of simulated people. (2) Run the model to let them age naturally until they reach the ages of the people who were candidates for the trial. (3) Identify those who would meet the entry criteria for the trial and select from them a sample that corresponds to the sample size of the real trial. (4) Randomize the simulated participants into groups and have simulated providers give them the treatments described in the trial. (5) Run the model for the duration of the trial. (6) Count the biological and health outcomes of interest that occurred to the participants in the simulated trial and compare them to the results observed in the real trial. A full validation requires several such exercises. For each exercise, we look at three things: “matching,” “span,” and “independence.” We also look at the “comprehensiveness” of the full set of validation exercises.

For matching, we ask how closely the results of the simulated study match those of the real study. Because it is a true simulation, the simulated results will be affected by sample size in an identical way as the results of the real trial. We say there is “statistical matching” when the simulated results match the real results within appropriate confidence intervals. Because smaller trials have wider confidence intervals, they are easier to match. We therefore seek out trials that have sufficient size to provide a meaningful test and ask that the match be within one standard deviation, not the traditional 95% confidence interval. We can also increase the

number of people in the simulation to tighten the confidence intervals for the simulation's results.

“Span” refers to the number of equations the validation covers. We consider validations that span multiple equations to be disproportionately more important than one-equation validations because they test not only individual pieces of the model but all the possible connections between them. Simulation of a typical clinical trial can involve scores of equations.

For “independence”, we identify three main levels of validation. We call the exercise a “two-star” validation when the results being compared were used to fit one or more of the equations. For a “three-star” validation, the exercise should compare outcomes in a study that were *not* used to fit the equations. For example, we might use the effect of a treatment on LDL-cholesterol after one month to set up a validation and then use the rate of strokes at 10 years to validate it. Such an exercise can be considered a “partial prediction” in the sense that the outcomes being calculated were never used by the model; long-term outcomes might be predicted from short-term outcomes or health outcomes might be predicted from biological outcomes. Finally, the model can simulate trials whose results were not used at all to build the model. This would represent a “full prediction” and we call it a “four-star” validation.

Finally, “comprehensiveness” refers to the portion of the total model covered by the full set of validation exercises to which the model was subjected. For a fully comprehensive validation of an application, every equation involved in calculating an outcome of interest should be covered by at least one validation exercise. This approach to validation implies that a model can be validated for certain applications (e.g., calculating the effect of a particular treatment on a particular outcome), but not validated for a different application (e.g., the effect of a different treatment that is less well understood). We do not consider the model to be “valid” in any general sense, as though it were a property of the model for all possible applications.

16. Discussion

This model differs from other models in health care in several ways. First, it is an object-oriented simulation. All of the objects considered by clinicians and administrators to be important are represented as objects in the model, one-to-one. Second, it is written at a high level of anatomic, biologic, and clinical detail—roughly the level of detail at which clinical textbooks are written and research is designed. Unlike other models that include biological variables as inputs, an Archimedes model is continuously calculating all the pertinent biological variables and their interactions as part of a simulated physiology. This enables the model to analyze

detailed practice guidelines, disease management programs, and performance measures. It also enables the model to address combinations of interventions and comorbidities in a natural way. Finally, the biological grounding of the model makes it easy to revise when new information becomes available or new interventions are introduced. A third difference is its breadth. The model includes not only diseases and the clinical interventions that affect them, but also the other elements of a health care system such as care processes, facilities, personnel, equipment, and budgets. These aspects of the model are also written at a high level of detail, as determined by the needs of clinicians and administrators.

Mathematically, the model is distinguished by the use of differential equations. This has two important effects. One is to preserve the continuous nature of biological variables. For example, fasting plasma glucose (FPG) is not modeled in discrete states such as “normal,” “prediabetes,” and “diabetes.” Rather, the FPG of every person is calculated continuously as part of a person's physiology and can be measured at any instant. All the biological variables and interactions pertinent to the clinical management of a problem are calculated in this way. The use of differential equations also makes the model completely continuous in time; any event can occur at any moment. This is important because many decisions in health care involve timing. Examples include how frequently to screen a person or monitor a patient; when to initiate, modify or switch a treatment strategy; how frequently to schedule follow-up visits; the dependence of the sensitivity of a test on the degree of development of a disease (e.g., size of a mass); and the effectiveness of a treatment on the time it is given (e.g., thrombolytics for myocardial infarctions).

A final distinction is the method of validation. The fact that Archimedes is a detailed object-oriented simulation model provides a rigorous way to conduct validations; the model can be asked to reproduce or predict the results of real clinical trials. These simulations test dozens of equations and their interactions. Simulations of several trials can criss-cross different parts of the model in ways that make successful matches exceedingly unlikely to be due to chance, and that virtually ensure that the underlying formulation of the physiology is accurate.

The level of detail of the model is determined by three things. The first is the level of detail of the clinical and administrative questions being asked. For example, if the intention is to analyze a disease management program for congestive heart failure, and one of the variables clinicians are debating is the degree of cardiac failure that should be used to assign patients to a case manager (e.g., ejection fractions less than 50% versus ejection fractions less than 40%), then the model has to include cardiac output and ejection fraction, and all the

other variables that determine them and might be affected by treatments (e.g., myocardial contractility). The second determinant is the level of detail clinicians require to have confidence that the model includes all the relevant factors. In general, this means that the model will include all the variables they use to determine the management of a patient. The third determinant of the model's detail is the availability of research studies that provide empirical information about the variables.

The last criterion means that the scope, depth, and uses of the model will ultimately depend on the quantity and quality of the available studies. Because the model is written at the level of physiological detail that matches the available research, there is a straightforward approach to finding information about a variable: When experts consider a variable to be important, it is almost always because there is research that demonstrates its importance; that research can be used to include the variable in the model. A corollary of this is that if there is a well-designed study that addresses a particular set of variables, then the results of that study can be used to include those variables. Conversely, if there are no good studies that address a particular variable or set of variables, then the model will not include those variables for use in a predictive fashion.

As with any model, if there are important omissions or errors in a dataset, if there are biases or omissions in the conduct or reporting of a study, the equations will be thrown off accordingly. Furthermore, because the equations for different parts of the model can come from different sources and because of the possibility that the different sources might be from different populations and/or use different methods, there is always a risk of producing a camel—a model that might be correct in its parts, but is not accurate in the whole. The best way to test this is through the validation exercises.

The types of simulations that can be performed by the model provide very deep validations. Even a two-star validation is much more difficult and meaningful than for other types of models because of the large number of equations that are required to make the physiologies of the agents function accurately enough to just survive the simulation, much less to match the real results accurately. For example, simulation of an epidemiological study of the incidence of a disease requires that about a dozen pieces of the model all function properly, not just as separate parts, but in a connected way. They include the equations that govern the progression of the primary features that cause the disease as functions of all the pertinent risk factors; the occurrence of symptoms as a function of the primary features; the behavior of patients in recognizing and seeking care for symptoms; the making of appointments for the symptoms; the ordering of tests by simulated physicians; the accuracies of the tests; the interpretation of the tests to make the diagnosis; and the accurate calculation of incidence rates by

a simulated epidemiologist. Simulations of clinical trials that involve several different treatment protocols (e.g., “treat LDL-cholesterol to a goal of 130 mg/dL”) add higher orders of difficulty.

Three- and four-star validations provide even greater confidence that the model's representation of human physiology and disease is accurate. However, the “failure” of a three- or even four-star validation does not necessarily mean that the model was wrong. One reason is that different trials that appear to be addressing the same question can produce different results. The model will be just as confused by this as are experts. A second reason is that trials often produce surprising results. A four-star validation means that the results of the trial were completely predictable from existing information (predictable by the model at least). If a trial's results are surprising to experts, they will be surprising to the model. It should also go without saying that the model cannot predict events that have never been observed. For example, the model would not have predicted that fen-phen (fenfluramine, phentermine, and dexfenfluramine) would affect mitral valves.

With these limitations in mind, when new or surprising results occur, they can be incorporated into the model at the level of biologic detail at which the mechanisms of action are understood to occur. As more and more observations are incorporated into the model, the equations describing the underlying biology of diseases and their treatments become increasingly robust.

For parts of the model that are spanned by overlapping validations, the model can be used for a variety of purposes. They include: (1) combining the results of existing research (e.g., comparing two treatments that have been well researched separately, but never compared directly); (2) analyzing combinations of treatments (e.g., treat LDL cholesterol, blood pressure, and weight); (3) analyzing issues of timing; (4) extending the results of trials into the future (e.g., calculating long-term health outcomes from short-term biological outcomes); (5) extending research to different populations (e.g., patients with more severe diseases, or with different combinations of risk factors); (6) analyzing the effects of changes in care processes (e.g. disease management programs, continuous quality improvement programs); (7) analyzing the effects of clinical interventions and care processes on logistics and costs; (8) trying out different guidelines and disease management programs to improve their designs and suggest short-term outcomes that should be monitored to confirm their effects; (9) identifying the most important determinants of an outcome; (10) developing quantitative rankings of different interventions across diseases, for setting priorities and designing performance measures; (11) determining the extent to which the results of a program observed in one setting are transferable to other settings and the factors

that determine the transferability; (12) conducting cost-effectiveness analyses; and (13) planning research.

Although it is not the subject of this paper, a note about the modeling of care processes is appropriate. In practice, we develop our models of care processes through examination of existing protocols, interviews, and on-site observations, checked against any available data. Pilot studies can be conducted as needed. Many of the data required to model system resources have already been collected for administrative purposes. However, some are more problematic. An example is how much money will be saved if a resource is not used (e.g., a bed goes empty because a myocardial infarction is prevented, but no nurses are laid off). In such cases, administrators must describe how they would handle different scenarios and must understand that the accuracy of the results will depend on the accuracy of their descriptions. Any decision rules they might apply can be easily added to the model (by putting simulated administrators in the model and giving them a function of applying the decision rules).

As with any model, the key to addressing the data problem, whether it is on the clinical or logistic side of the model, is to appreciate that as the quality of the evidence varies the appropriate uses of the model change. At one extreme, if the parts of the model pertinent to a particular application can be well validated, then it is reasonable to act on the model's results. At the other extreme, if there is no information at all about certain variables, uses of the model that require those variables should be restricted to building understanding. The appropriate techniques here include "what if" calculations, identification of the most critical variables, identification of upper and lower bounds for an outcome, and planning of research. As with any model, sensitivity analysis can be used to explore the vulnerability of a conclusion to any particular variable or combination of variables.

Currently, the model includes the biological variables pertinent to coronary artery disease, congestive heart failure, diabetes, and asthma. About 6 months is needed to add a new disease. The model is programmed in Smalltalk (a pure object environment and programming language) and runs on any personal computer. A typical calculation for a medical center serving 300,000 people takes 24–72 hours. Computation times can be reduced by running parallel calculations.

Acknowledgments

Ken Heikes, Ph.D., Gerardo Soto-Campos, Ph.D., and Leonid Malozemov, Ph.D., read and reviewed the original manuscripts and made many helpful suggestions. Leonid Malozemov, Ph.D., helped in the description of random processes. Jon Wright, Ph.D., suggested we look into the KL decomposition and provided several references. The Medical Editing Department of Kaiser Foundation Hospitals, Inc., provided editorial assistance.

References

- [1] Karhunen K. *Über lineare methoden in der wahrscheinlichkeitrechnung*. *Am Acad Sci, Fennicae, Ser A*, 1 1947;37:3–79, Translation: RAND Corporation. Santa Monica (CA): RAND Corp.; 1960 (Report T-131).
- [2] Loeve M. *Fonctions aleatoires du second ordre*. In: Levy P, editor. *Processus stochastiques et mouvement brownien: suivi d'une note de M Loeve*. Paris: Gauthier-Villars; 1948.
- [3] Kac M, Siebert AJ. An explicit representation of a stationary gaussian process. *Ann Math Statist* 1947;18:438–42.
- [4] Ghanem RG, Spanos PD. *Stochastic finite elements: a spectral approach*. 1991 New York: Springer; 1991.
- [5] Newman AJ. *Model reduction via the Karhunen–Loeve expansion. Part I: An exposition*. [College Park (MD)]: Institute for Systems Research, University of Maryland; 1996. Technical Research Report, T.R. 96-32. Available on the World Wide Web (accessed June 11, 2001): http://www.isr.umd.edu/TechReports/ISR/1996/TR_96-32/TR_96-32.phtml.
- [6] StatSoft, Inc. 2001. *Electronic statistics textbook*. Tulsa (OK): StatSoft; 2001. Available on the World Wide Web (accessed May 11, 2001): <http://www.statsoft.com/textbook/stathome.html>.
- [7] Bock RK, Krischer WK. *The data analysis brief book* [Internet ed.]. Geneva: European Organization for Nuclear Research; [1999]. Available on the World Wide Web (Accessed May 11, 2001): <http://rkb.home.cern.ch/rkb/titleA.html>.
- [8] United States. National Heart, Lung, and Blood Institute. Framingham Heart Study: 50 years of research success [website]. Available on the World Wide Web (accessed May 11, 2001): <http://www.nhlbi.nih.gov/about/framingham/>.
- [9] Roberts WC. Qualitative and quantitative comparison of amounts of narrowing by atherosclerotic plaques in the major epicardial coronary arteries at necropsy in sudden coronary death, transmural acute myocardial infarction, transmural healed myocardial infarction and unstable angina pectoris. *Am J Cardiol* 1989;64(5):324–8.
- [10] Selwyn AP, Braunwald E. Ischemic heart disease. In: Isselbacher KJ, editor. *Harrison's principles of internal medicine*. 13th ed. New York: McGraw-Hill, Health Professions Division; 1994. p. 1077–8.