25
MODELS IN MEDICINE

Michael Wilde and Jon Williamson

1. Introduction

The major goals of medicine include predicting disease, controlling disease, and explaining disease. The main way of achieving these goals proceeds by modelling. In this chapter, we provide an introduction to the use of models in achieving the goals of medicine. To begin with, we introduce the notion of a model in medicine and distinguish experimental models from theoretical models. Then we provide an overview of the extensive array of these models by giving an account of animal models, which are a kind of experimental model, as well as association models, causal models, and mechanistic models, which are kinds of theoretical models. Next, we argue that in order to achieve the goals of medicine we need all of the latter three kinds of theoretical model—none are redundant. We go on to provide a framework for systematizing the production of theoretical models. Lastly, we present an example involving benzene and leukemia to illustrate the approach of this chapter.

2. Models in Medicine

The use of models is an important feature of scientific practice. Accordingly, scientific models have received a good deal of attention from philosophers of science. This attention has tended to focus on general problems such as the nature of models, how models are related to scientific theories, and how scientists can learn about the world by using models (Frigg and Hartmann, 2005). The scope of this chapter is much narrower, concerning only models in medicine. Given this, we shall set aside a number of important general issues in order to focus on those matters most pertinent to medicine.

Broadly speaking, we take a model to be a structure that represents some target system and that is used as a means of drawing conclusions about that target system. It is difficult to draw conclusions directly about a target system where that target system is inaccessible or very complicated. Usually, in such cases, it is more straightforward to reason instead about a model, since the model may involve considerable simplifications. The utility of models lies in the fact that conclusions drawn from the model can carry over to the target system, as long as the model is a sufficiently good representation of that target system.

As far as models in medicine are concerned, the target system is typically an individual or population of biomedical interest. The aim is to draw conclusions about such an individual or population. The conclusions that are relevant to medicine include claims about associations and causal relationships between exposures and diseases, as well as claims about biological mechanisms. On the one hand, it is important to establish associations and causal relationships in medicine, since the goals of medicine include predicting and controlling disease,

and it is not possible to achieve these goals without information about associations or causal relationships (Russo and Williamson, 2007). On the other hand, it is also important to find out about biological mechanisms, since another of the goals of medicine is to explain disease, and it seems that explanations are best given by appealing to mechanisms (Williamson, 2013).

There have been a number of attempts to classify the different types of model in science, but none of them seems entirely satisfactory (Mäki, 2001). However, for our purposes, models in medicine can be usefully classified into two types, experimental models and theoretical models.

An experimental model is typically a concrete object that is experimented upon in order to draw conclusions about the model. This experimentation also licenses conclusions about a target system, insofar as the experimental model adequately represents the target system. The experimentation is intended to gather brand-new information about the target system. A theoretical model is a more abstract construct that systematizes information that has already been gathered about the target system. This systematization allows further conclusions to be drawn from that information more straightforwardly than would be possible if the information were not systematized in a model.

We shall shortly fill in the details of this classification by presenting animal models as examples of experimental models, and then presenting association models, causal models, and mechanistic models as examples of theoretical models. The aim is to show exactly how all of these models help achieve the major goals of medicine, i.e., predicting and explaining the occurrence of diseases, as well as providing recommendations about the control of such diseases.

3. Animal Models

One kind of model is an *animal model*, e.g., an *experimental organism*. An experimental organism, at least as far as medicine is concerned, is a non-human organism that is experimented upon in order to gather information relevant to the prediction, explanation, and control of disease in humans. A particular experimental organism is typically chosen on the grounds of its tractability to experimentation and its suitability to the biomedical phenomenon under investigation. Often, practical considerations, such as the availability of the organism for investigation, also inform the selection of experimental organism (Kohler, 1994).

An experimental organism is a means of gathering biomedical information about humans in cases where it is not possible to gather this information by directly conducting experiments involving human subjects, e.g., where such experiments would be unethical or difficult to carry out. As long as the experimental organism is representative of humans in the appropriate respects, conclusions arrived at by experimenting upon the organism also license further conclusions about humans. The conclusions may include claims about associations between exposures and disease, as well as claims about biological mechanisms.

A famous example of a model from the history of experimental physiology is the frog, which was studied in order to learn about the biological mechanisms of muscle contraction in humans and mammals more generally (Holmes,1993). In this case, it was difficult to learn about muscle contraction in humans directly, since it was not possible to carry out the relevant experiments on humans. Instead, the frog was experimented upon as a representative model of the mechanisms of human muscle contraction. The conclusions drawn about muscle contraction in frogs were taken to apply also to muscle contraction in humans, insofar as the frog was appropriately representative of human muscle contraction.

A further example is the use of experimental organisms in preclinical animal trials for toxicology testing. Such tests are often conducted using mice or rats in order to assess the safety

or efficacy of a drug before comparative clinical trials in humans are attempted. A comparative clinical trial in humans is carried out only if it has been established that the drug is not associated with adverse outcomes in the experimental organism. This is because establishing the safety of the drug in the experimental organism is taken to support the conclusion that the drug is safe also in humans, to a sufficient extent that it is deemed safe for comparative clinical trials in humans to proceed.

As these examples make clear, the use of experimental organisms can help with all of the main goals of medicine. First, experimental organism research can help with explanation because it allow claims about biological mechanisms in humans to be established (as long as the experimental organisms are appropriately representative). Second, it can help with predicting and controlling disease since it enables claims about associations and causal relationships between exposures and disease to be established (again, as long as the organisms are representative).

There have been a number of debates about experimental organism research. Claude Bernard (1865) believed that the results of animal experiments were straightforwardly applicable to humans, since the differences between animals and humans were only a matter of degree. However, Hugh LaFollette and Niall Shanks (1996) have argued that evolutionary theory casts doubt on the claim that it is justified to extrapolate from experimental organisms to humans, and that this makes experimental organism research morally questionable. Recently, it has been argued that significant findings in preclinical animal trials rarely lead to successful treatments in humans (Djulbegovic et al., 2014). Some have suggested that this may be because many animal trials are poorly conducted (Hirst et al., 2014).

Rachel Ankeny and Sabina Leonelli (2011) have argued that *model organisms* should be distinguished from the broader class of experimental organisms. Some examples of model organisms include the fruit fly, the nematode worm, and certain strains of mouse. Among other things, Ankeny and Leonelli argue that model organism research is unlike experimental organism research in that it aims to provide a detailed account of the model as a whole organism, in terms of its genetics, physiology, evolution, and so on. Arnon Levy and Adrian Currie (2015) have argued that model organisms are not models in the traditional sense. In traditional modelling, conclusions about the target system are supported by assessing whether the model is sufficiently analogous to the target system. In model organism research, they argue, the models are not analogues of some target system but instead are samples from a broader class of organisms. They maintain that the conclusions drawn from model organisms are the result of empirical extrapolation mediated by indirect evidence concerning the similarity of members of a broader class, where this broader class includes both the model organism and its target system. This indirect evidence might be that the broader class of organisms have a shared evolutionary ancestry or shared phylogeny. Marcel Weber (2005) argues that extrapolations from model organisms to humans can be reasonably sound, as long as they are based on known phylogenetic relationships.

This concludes our discussion of animal models. Now we survey the principal kinds of theoretical model: association models, causal models, and mechanistic models.

4. Association Models

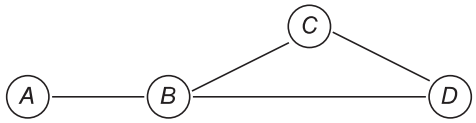
One simple kind of theoretical model employed in medicine is an association model. This charts the main correlations among variables measured in a data set, so that one can use the observed values of certain measured variables to predict the value of an unmeasured variable in a new patient.

When applied to diagnosis, for example, an association model might be used to determine the probabilities of a range of possible diseases, given a particular combination of symptoms observed in a particular patient. These probabilities can then be used to motivate a particular treatment

decision. An association model for prognosis, on the other hand, will usually be used to predict severity of disease, given observed clinical features of the patient and any observed biomarkers of the disease in question. Either way, the main use of the association model is *prediction*.

By way of example, we shall present two kinds of association model: a Markov network model and a Bayesian network model.

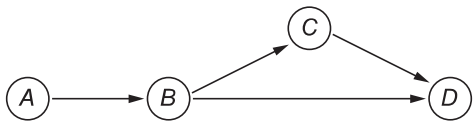
From a qualitative point of view, an association model can typically be represented by an undirected graph, sometimes called a *Markov network* , with nodes corresponding to variables and edges corresponding to the significant associations:



Separation in the graph can be used to denote probabilistic independence. In the above graph, B separates A from C and D, in the sense that all paths from A to C or D proceed via B. This separation relationship can be used to signify that A is probabilistically independent of C and D, conditional on B. (A is probabilistically independent of C and D, conditional on B, written $A \perp\!\!\!\perp C, D \mid B$, just when $P(a \mid bcd) = P(a \mid b)$ for all values a, b, c, d, of A, B, C, D respectively.) Thus, if one wants to predict A and one can observe B, it would make no sense to also observe C and D, because these would provide no further information about A. For example, suppose that a blockage in the main coronary artery (A) raises the probability of a heart attack (B), which in turn raises the probability of particular electrocardiogram results (C and D), in such a way that can be charted by the above association model. Then, to predict that the patient has had a blockage in the main coronary artery, one need only observe that the patient has had a heart attack, since learning in addition that the patient had certain electrocardiogram results provides no more information about the blockage.

From a quantitative point of view, in order to determine the probability of any variable conditional on any given combination of values of the other variables, one needs to specify the joint probability distribution, defined over all the variables of interest. In the above example, one would need to specify $P(abcd)$ for each combination of values a, b, c, d, of A, B, C, D, respectively. In a Markov network this is achieved by specifying the probability distribution over variables in each clique of the graph. A clique is a maximal subset of nodes of the graph such that each pair of variables in the subset is connected by an edge. The cliques in the above graph are $\{A, B\}, \{B, C, D\}$, so one would need to specify $P(ab)$ and $P(bcd)$ for all combinations of values a, b, c, d, of A, B, C, D, respectively.

Alternatively, one can use a *Bayesian network* model to represent the joint probability distribution. A Bayesian network has a qualitative and a quantitative component. The qualitative component of a Bayesian network consists of a directed acyclic graph, i.e., a graph with arrows such that there is no path in the direction of those arrows from a node to itself. In our example, one possible directed acyclic graph would be:



The directed acyclic graph needs to be constructed in such a way that each variable is probabilistically independent of its non-descendants, conditional on its parents. (A non-descendant of a variable is any node that cannot be reached by a directed path from the variable in

question. A parent of a variable is any node from which there is an arrow to the variable in question. For example, in the graph below, *A* is a parent of *B*, and *B* is a parent of C and *D*. This means that C and *D* are descendants of *B*, and so *B* is a non-descendent of both C and *D*.) The quantitative component of a Bayesian network consists of the probability distribution of each variable conditional on its parents. The probability of a particular combination of values of variables is then a product of specified conditional probabilities:

$$P(abcd) = P(a)\ P(b \mid a)\ P(c \mid b)\ P(d \mid cb).$$

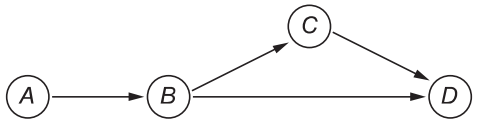
There are a wide range of algorithms for producing a Bayesian network that represents the observed probability distribution of a set of variables measured in a data set (see, e.g., Neapolitan, 2004). There are also many algorithms for drawing predictions from a Bayesian network (see, e.g., Darwiche, 2009). Note that the directions of the arrows in the Bayesian network do not represent causal relationships in this sort of association model—the arrows are merely a technical device for representing certain probabilistic independencies.

Another kind of association model, called a classifier, is often used when it is only necessary to predict the value of a single variable, such as severity of disease. Many such models have been devised in the fields of machine learning and statistics (see, e.g., Alpaydın, 2010).

5. Causal Models

A second kind of theoretical model widely used in medicine is a causal model. A causal model seeks to chart the causal connections between the variables of interest. Such a model has three uses: prediction, explanation, and control. Like an association model, a causal model can be used for prediction, since it can be used to infer the probability of one variable conditional on others taking certain observed values. It can also be used to construct rudimentary explanations, since one can explain the fact that a particular variable takes the value that it does in terms of the causes of the variable in question taking certain values. Most importantly, perhaps, it can be used for control: it can be used to predict the effects of interventions and so can be used to decide how best to intervene in order to control the disease or symptoms of a particular patient.

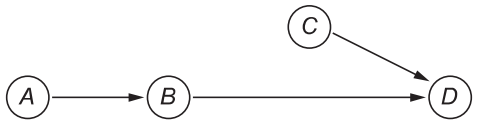
A causal model can represent causal connections qualitatively by means of a directed acyclic graph. For example:



In contrast to the arrows of the directed acyclic graph presented in the previous section, which featured in an association model, in a causal model the arrows have significance in that they represent direct causal connections. For example, the above graph says that A is a cause of C, but only via a single pathway that proceeds through *B*.

A *causal Bayesian network* or *causal network* is a Bayesian network model built around a causal graph such as the above. Formally, it is a Bayesian network, but now the arrows in the graph have causal significance. Because it is a Bayesian network, it can be used to define a joint probability distribution over the variables in the graph, and thus can be used for prediction. But because the arrows have causal significance, it can also be used to predict the effects of interventions, as follows (see Pearl, 2000). When an intervention is performed to fix a variable to a certain specific value, one modifies the Bayesian network by deleting all arrows into this variable in the graph and updating the conditional probability distribution

of each variable conditional on its parents in the graph to take into account the new value of the intervened-upon variable. Then this modified network can be used to infer changes to the probabilities of variables of interest, given the intervention. For example, intervening to fix the variable C to a specific value c will lead to a modified network in which the arrow from *B* to *C* is deleted:



There are other sorts of causal model besides causal Bayesian networks (Illari et al., 2011); such models tend to portray causal processes in a similarly schematic way, representable by means of a directed acyclic graph. The explanations offered by such models can be superficial in that they only pick out key variables—milestones on the causal pathways to the effect in question—rather than the detailed structure of the underlying mechanisms that are responsible for the phenomena to be explained.

6. Mechanistic Models

Mechanistic models are used to generate explanations that are less superficial than the explanations yielded by causal models, in that they tend to include a richer set of explanatory features. There are two principal sorts of mechanism. A *complex-systems mechanism* consists of entities and activities organized in such a way that they are responsible for some phenomenon of interest (Machamer et al., 2000; Illari and Williamson, 2012). Examples include the mechanism for the circulation of the blood (which includes the features responsible for operation of the heart as well as the organization of the other components of the cardiovascular system) and the mechanism in an artificial pacemaker for producing electrical impulses to stimulate the heart (which includes its power source, clock, sensors and pulse generator, and the features of their arrangement that ensure its correct operation). On the other hand, what one might call a *mechanistic process* is a spatio-temporally contiguous process along which a signal is propagated (Reichenbach, 1956; Salmon, 1998). Examples include the process of an artificial pacemaker's electrical signal being transmitted along a lead from the pacemaker itself to the appropriate part of the heart, and the process by which an airborne environmental pollutant reaches the lining of the lung. While complex-systems mechanisms are often multi-level—e.g., involving coordinated activity at the levels of the organism, the organ, the cell, and the gene—mechanistic processes usually take place at a single level. Furthermore, whereas complex-systems mechanisms typically operate in a regular way, repeatedly producing the phenomenon of interest, mechanistic processes are often one-off, transmitting a single signal on a single occasion. In either case, however, the mechanism's structure and its organization—particularly its spatio-temporal organization—tends to be crucial to its operation.

A mechanistic explanation will often appeal to both sorts of mechanism. An explanation of the circulation of the blood in a particular individual may appeal to the complex-systems mechanism by which the heart pumps the blood, as well as the complex-systems mechanism of the individual's pacemaker and the mechanistic process linking the two. An explanation of a failure of blood to circulate may appeal to the same mechanisms, any faults of these mechanisms, and any pathophysiological mechanistic processes that these faults give rise to.

Mechanistic models are used to model the salient features of mechanisms in order to explain phenomena of interest. They differ from causal models in that they appeal to a richer set of features—entities, activities, organization, hierarchical structure, processes, etc.—instead of

simply variables or events. Some of these features cannot be easily incorporated into a causal model: spatio-temporal organization and hierarchical structure, for example, are not naturally represented using the nodes and arrows that typically characterize causal models. However, these features are often essential components of an adequate explanation. Only in cases where these features are not essential to the explanation will an explanation generated from a causal model be adequate, in the sense that it picks out all the main features of an adequate mechanistic explanation (Williamson, 2013).

We noted above that a single mechanistic model may seek to represent two kinds of mechanism: complex-systems mechanisms and mechanistic processes. In addition, mechanistic models can be classified into two kinds: qualitative and quantitative.

Qualitative mechanistic models fill textbooks and research papers in medicine. They usually take the form of diagrams that highlight the main features of the mechanism. For example, Figure 25.1 portrays a part of the mechanism for apoptosis (cell death). Increasingly, animations are employed as qualitative mechanistic models, in order to portray activities and processes developing over time. Agent-based models are another kind of qualitative mechanistic model. Such a model represents a target system, e.g., a human population, in terms of a large numbers of similar individuals that interact in a restricted set of ways, e.g., colored cells in a grid that

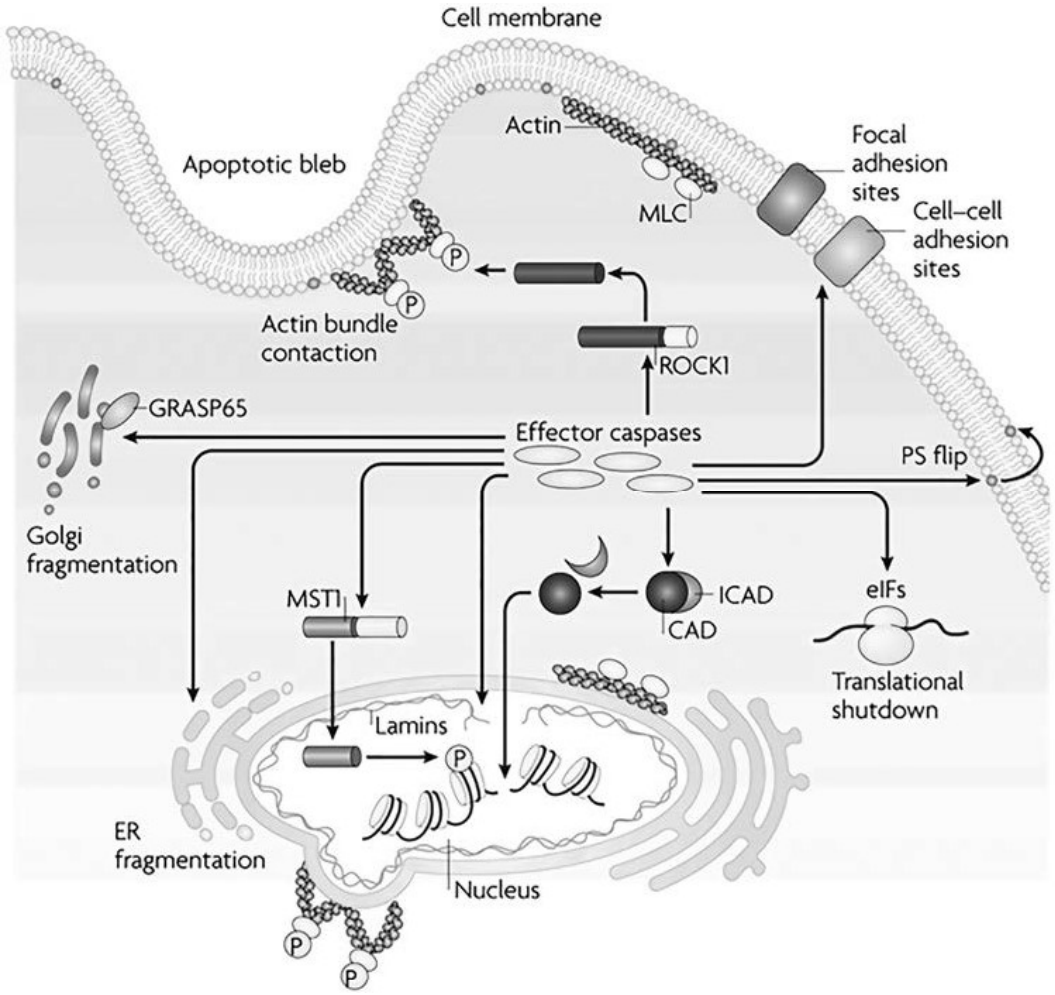


Figure 25.1 Caspases coordinate demolition of key cellular structures and organelles (Taylor et al., 2008)

influence the colors of their neighbours. Computer simulations are used to determine the typical behaviour of such a system. To the extent that this simulated behavior tallies with some observed phenomenon, such as the spread of a contagious disease, the agent-based model can be used to explain the occurrence of phenomenon.

Qualitative mechanistic models can be used to point to the underlying structure of reality that is responsible for producing the phenomenon to be explained, but normally cannot, on their own, explain why certain quantities within the mechanism take the values that they do, or explain the probability of a certain phenomenon. For this sort of explanation, a quantitative mechanistic model is required. A quantitative mechanistic model might, for example, consist of a diagram that portrays the qualitative structure of the mechanism, together with differential equations that can be used to model the changes in certain quantities over time. Another example of a quantitative mechanistic model is a *recursive Bayesian network*, which can represent a hierarchically structured mechanism by means of a collection of causal Bayesian networks, and which can be used to infer the probability of variables in the mechanism, given the observed values of other variables (Casini et al., 2011; Clarke et al., 2014b).

7. Combinations of Theoretical Models

Given this extensive array of theoretical models—association, causal, and mechanistic—two questions arise. Do we really need all these kinds of model in medicine? If so, is there any way of systematizing and unifying the production of these models? We will argue in this section that both of these questions should be answered affirmatively.

Do we need all these models? As we hope to have made clear in the above discussion, different kinds of theoretical model are put to different uses. Association models are for prediction; causal models are for prediction, explanation, and control; mechanistic models are primarily for explanation. One might think, then, that in medicine we should strive to produce good causal models, which can be put to the widest variety of uses, and we should avoid association and mechanistic models. There are three main reasons why this is not a sensible suggestion.

First, as we have mentioned, causal models generate more impoverished explanations than do mechanistic models. Causal models abstract away from the details of mechanistic structure, generating explanations that invoke only variables and the "thin" causing relation, i.e., explanations that invoke only claims of the form X causes Y. Mechanistic models, on the other hand, invoke entities, "thick" activities such as dilating and osmosing (i.e., a rich variety of kinds of causing), organizational features such as the structure and location of the cell wall, constitutive relations between components at different levels of a hierarchical mechanism, and spatio-temporally contiguous processes. Therefore, mechanistic models are far from redundant in situations in which a detailed explanation is required.

Second, causal and mechanistic models tend to be less reliable than association models. It is relatively easy to build an association model from some given data: one merely needs to model the joint probability distribution that generates the data. Often, in cases in which there are ample, good-quality data and no reason to suspect bias in the way the data were sampled, one simply models the joint distribution of the data and treats that data distribution as an approximation to the data-generating distribution. In a causal model, however, one needs to model not only the associations in the data but also the causal relationships among all the measured variables. Determining causal relationships is a harder problem than determining associations, so a causal model will normally be more speculative than an association model. Harder still is the task of establishing the details of the underlying mechanisms. This has long been the primary goal of biomedical science, and while great progress has been made, many mechanistic models are either very speculative or "gappy,"

with important features missing. This is less the case with a causal model: one needs to establish causal connections between those variables that are in the model, but there is no requirement to include in the model every variable that represents a component of one of the pertinent mechanisms. A causal model that omits some salient variables can still generate useful inferences for prediction and control, and capture some explanatory factors. In sum, there is a sense in which association models are normally less speculative than causal models, which are in turn normally less speculative than mechanistic models. Association models, in particular, retain an important place in medical research.

Third, association and mechanistic models are epistemically prior to causal models. This is a consequence of the following epistemological thesis, put forward by Russo and Williamson (2007). In order to establish a causal claim in medicine, one normally needs to establish two things: (1) that the putative cause and putative effect are appropriately correlated; and (2) that there is some underlying mechanism that links the cause to the effect in an appropriate way and that explains this correlation.

Some points of clarification are needed. First, the latter two claims are existence claims: in order to establish causality, one normally needs to establish the existence of a correlation and the existence of a mechanism, not the precise extent of the correlation nor all the details of the mechanism (Darby and Williamson, 2011, 2). Second, the mechanism involved might be a complex-systems mechanism, or a mechanistic process, or a combination of the two—whatever connects the putative cause to the putative effect in such a way that can *explain* occurrences of the latter. Third, this thesis concerns the evidence required to *establish* a causal claim, i.e., to settle the question according to the standards of the community, in such a way that warrants a high degree of confidence that the causal claim will not be overturned by any new evidence.

This epistemological thesis is plausible for the following reason. Recall that in medicine, causal claims are used for prediction, explanation, and control. If the putative cause and putative effect were not appropriately correlated, one would not be predictive of the other and one would not be able to intervene upon the cause to control the effect. Moreover, if there were not some mechanism that links the cause to the effect in an appropriate way, one would not be able to invoke the cause to explain the effect. Why is establishing a correlation not normally sufficient on its own for establishing causality? This is because many correlations are best explained by relationships other than causal connection—such as semantic, logical, or mathematical relationships—or by confounding, bias, or chance. Mechanistic evidence steers the causal discovery process toward those connections that are genuinely causal (Clarke et al., 2014a). Investigations of cases of causal discovery that provide further evidence in favor of the epistemological thesis include Clarke (2011), Darby and Williamson (2011), Gillies (2011), and Russo and Williamson (2011, 2012).

This epistemological thesis applies to each causal claim in a causal model. One therefore normally needs to establish the pattern of correlations, as represented by an association model, as well as the pattern of mechanistic connections, as represented by a mechanistic model, in order to establish the qualitative causal connections posited by the causal model. Association and mechanistic models are epistemically prior to causal models, since one needs to establish features of the former two kinds of model in order to establish the claims made by the latter kind. (Since the epistemological thesis merely requires one to establish existence of a correlation and a mechanism for each causal connection, in order to determine the pattern of causal relationships one only needs to establish the *pattern* of correlations and the *pattern* of mechanisms, not other features of association and mechanistic models.)

For these three reasons, one should not seek to abandon association and mechanistic models in favor of causal models.

How can the production of models be systematized? The third of the above three reasons suggests one way of systematizing and unifying the production of these models. First, consider an idealized case in which the available evidence is so extensive and of such high quality that it allows one to establish the full pattern of associations posited by an association model and the full pattern of mechanisms posited by a qualitative mechanistic model. Then one is in a position to establish the full pattern of causal claims made by a causal model, as well as the quantitative component of the causal model, which determines the joint probability distribution over the variables in the model. Having specified the quantitative component of a causal model, one is then in a better position to augment a qualitative mechanistic model by adding quantitative information.

Of course, in practice it is almost never the case that evidence is so plentiful and of such high quality as to establish every association and every mechanistic connection. In practice, evidence is often inconsistent, some data sets are extensive, others not, and items of evidence are of very varying quality. Thus some intermediate steps are needed to evaluate the relative merits of the items of evidence, and to determine which claims of association and mechanism can be considered established and which others are merely plausible or conjectural. It is possible, then, to establish some causal claims on the basis of what can be established in an association model and a mechanistic model. Other causal claims in the causal model will be more tentative, in proportion to the uncertainty of the corresponding association and mechanism claims.

This epistemological picture is depicted in Figure 25.2. Evidence of correlation (of which data sets are key) needs to be evaluated and graded with regard to the support it provides for associations. For example, data sets arising from larger numbers of observations will normally be more highly graded, and experimental studies will normally be favored over observational studies. Evidence of mechanisms informs this evaluation process because such evidence is crucial to determining whether trials are well-designed and their results correctly interpreted (Clarke et al., 2014a). This will lead to an association model. On the other hand, evidence

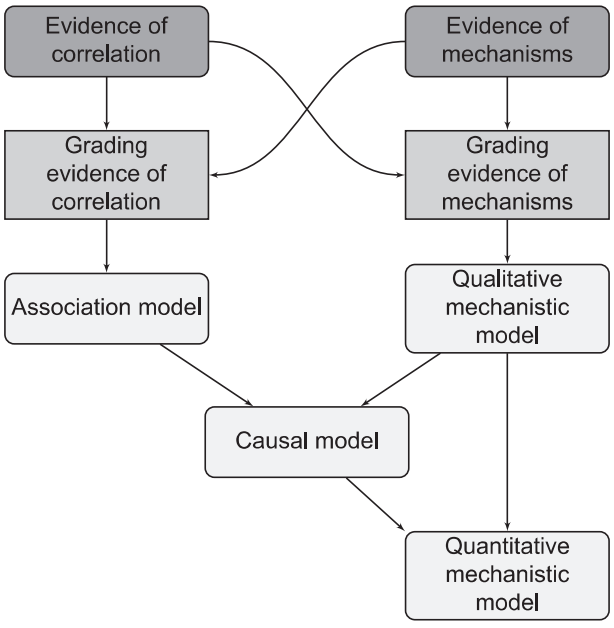


Figure 25.2 Relationships among theoretical models

of mechanisms (which can also be gained from basic lab research, imaging, autopsy, etc.) will need to be evaluated in order to construct a qualitative mechanistic model. Here, evidence of correlation is important in identifying the most salient components in the mechanisms and in determining the net effect of several interacting mechanisms or components of a mechanism. With an association model and a qualitative mechanistic model in place, one is in a position to construct a causal model and to determine whether each claim made by the model can be considered established or more conjectural. The quantitative causal model will go on to inform a quantitative mechanistic model.

The Bayesian network family of models can be used as a unifying formal framework that dovetails with this epistemological picture. As discussed above, a standard Bayesian network can be used as an association model. One way of constructing a Bayesian network from a range of data sets is provided by the *objective Bayesian network* approach: an objective Bayesian network represents the probability distribution that best fits the range of data available, where this optimal distribution is determined by the principles of objective Bayesianism (Williamson, 2005; Nagl et al., 2008). Next, a causal Bayesian network can be constructed from the Bayesian network association model and a qualitative mechanistic model. Finally, a recursive Bayesian network might be employed as a quantitative mechanistic model.

We should conclude this section by noting that the unified account presented above is not the standard way to approach the problem. Where causal Bayesian networks are advocated, it is usually in the context of a data mining approach: the idea is to learn a causal Bayesian network directly from data, in a similar way to the way in which association models are often constructed directly from data (e.g., Spirtes et al., 1993). In contrast, we advocate developing a qualitative mechanistic model on the way to producing a causal model. This is because we hold that causal relationships track mechanistic connections as well as associations, and because one needs to establish that a posited causal connection does indeed track these two things before one can consider the causal claim to be established.

8. An Example: Benzene and Leukemia

One example that illustrates the approach of this chapter concerns benzene and leukemia. Benzene is a clear and highly flammable liquid. Among other things, benzene is added to gasoline in order to reduce engine knocking, and it is also used in the manufacture of organic chemicals. A number of studies established a relationship between benzene exposure and leukemia in humans (Infante et al., 1977; Rinsky et al., 1981). These results are corroborated by studies in mice and other rodents (Goldstein et al., 1982; Cronkite et al., 1984). The association between benzene exposure and leukemia can be charted in an association model, which can be used to predict the disease given the environmental exposure. The relationship may even be charted in a causal model, e.g., a causal Bayesian network, insofar as the relationship is causal. But such a causal model could not yet provide anything other than an impoverished explanation, since it claims only that there exists a mechanism linking benzene exposure and leukemia, rather than providing the details of that mechanism.

This is an example of a more general problem in epidemiology, which is the study of health and disease in defined populations. A key working hypothesis in epidemiology is that diseases are often the result of environmental exposures. However, despite much epidemiological research, the biological processes linking many environmental exposures and diseases remain unknown. This is the case despite the technological advances in measuring certain biomarkers, i.e., biological markers of events at the molecular and physiological levels. Unfortunately, the details of these biological processes are required in order to provide less-impoverished explanations of the occurrence of disease.

Molecular epidemiology is a response to this state of affairs (Schulte and Perera, 1993). Molecular epidemiology is a branch of epidemiology that uses advances in biomarker technology in order to elucidate the biological mechanisms between environmental exposures and diseases. An important methodology in molecular epidemiology involves utilizing complementary studies in order to validate biomarkers that mediate between environmental exposures and disease outcomes (Vineis and Perera, 2007). For example, some studies may provide information associating a certain biomarker to a particular environmental exposure. Other studies may provide information relating a disease outcome to the same biomarker. By bringing together the results of these studies, the disease may be associated with the environmental exposure while providing some insight into the biological processes responsible for this association by highlighting the intermediate biomarkers (Russo and Williamson, 2012).

In the case of benzene and leukemia, studies revealed that certain chromosome aberrations were predictive of cancer in humans (Bonassi et al., 2000). In other case-control studies, those same chromosome aberrations were seen to be more frequently present in leukemia patients who had been exposed to benzene (Zhang et al., 2007). These results are corroborated by animal models (see, e.g., Eastmond et al., 2001). Not only, then, was a chain of associations established between benzene and leukemia, but also some insight was provided into the biological mechanism underlying this chain, i.e., the role of chromosomal aberrations; Vineis and Perera, 2007. These insights can be represented in a mechanistic model, and the model may be used to provide a less-impoverished explanation of leukemia in terms of exposure to benzene. Furthermore, details of the mechanism underlying the association between benzene exposure and leukemia, along with the details of the association, can all be charted in a quantitative mechanistic model.

References

Alpaydın, E. (2010). *Introduction to machine learning*, second edition. MIT Press, Cambridge, MA.

Ankeny, R., and Leonelli, S. (2011). What's so special about model organisms? *Studies in History and Philosophy of Science*, 42:313–323.

Bernard, C. (1865). *An Introduction to the study of experimental medicine (1949)*. Henry Schuman, New York.

Bonassi, S., Hagmar, L., Strömberg, U., Montagud, A. H., Tinnerberg, H., Forni, A., Heikkilä, P., Wanders, S., Wilhardt, P., Hansteen, I.-L., Knudsen, L. E., and Norppa, H. (2000). Chromosomal aberrations in lymphocytes predict human cancer independently of exposure to carcinogens. *Cancer Research*, 60(6):1619–1625.

Casini, L., Illari, P. M., Russo, F., and Williamson, J. (2011). Models for prediction, explanation and control: Recursive Bayesian networks. *Theoria*, 26(1):5–33.

Clarke, B. (2011). *Causality in medicine with particular reference to the viral causation of cancers*. PhD thesis, Department of Science and Technology Studies, University College London, London.

Clarke, B., Gillies, D., Illari, P., Russo, F., and Williamson, J. (2014a). Mechanisms and the evidence hierarchy. *Topoi*, 33(2):339–360.

Clarke, B., Leuridan, B., and Williamson, J. (2014b). Modelling mechanisms with causal cycles. *Synthese*, 191(8):1651–1681.

Cronkite, E., Bullis, J., Inoue, T., and Drew, R. (1984). Benzene inhalation produces leukemia in mice. *Toxicology and Applied Pharmacology*, 75(2):358–361.

Darby, G., and Williamson, J. (2011). Imaging technology and the philosophy of causality. *Philosophy & Technology*, 24(2):115–136.

Darwiche, A. (2009). *Modeling and reasoning with Bayesian networks*. Cambridge University Press, New York.

Djulbegovic, B., Hozo, I., and Ioannidis, J. P. A. (2014). Improving the drug development process: More not less randomized trials. *Journal of the American Medical Association*, 311(4):355–356.

Eastmond, D., Schuler, M., Frantz, C., Chen, H., Parks, R., Wang, L., and Hasegawa, L. (2001). Characterization and mechanisms of chromosomal alterations induced by benzene in mice and humans. *Research Report, Health Effects Institute*, 103:1–68.

Frigg, R., and Hartmann, S. (2005). Scientific models. In Sarkar, S., editor, *The philosophy of science: An encyclopedia*, pages 740–749. Taylor & Francis, New York.

Gillies, D. A. (2011). The Russo-Williamson thesis and the question of whether smoking causes heart disease. In Illari, P. M., Russo, F., and Williamson, J., editors, *Causality in the sciences*, pages 110–125. Oxford University Press, Oxford.

Goldstein, B., Snyder, C., Laskin, S., Bromberg, I., Albert, R., and Nelson, N. (1982). Myelogenous leukemia in rodents inhaling benzene. *Toxicology Letters*, 13(3–4):169–173.

Hirst, J. A., Howick, J., Aronson, J. K., Roberts, N., Perera, R., Koshiaris, C., and Heneghan, C. (2014). The need for randomization in animal trials: An overview of systematic reviews. *PLoS ONE*, 9(6):1–11.

Holmes, F. (1993). The old martyr of science: The frog in experimental physiology. *Journal of the History of Biology*, 26:311–328.

Illari, P. M., Russo, F., and Williamson, J., editors. (2011). *Causality in the sciences*. Oxford University Press, Oxford.

Illari, P. M., and Williamson, J. (2012). What is a mechanism? Thinking about mechanisms *across* the sciences. *European Journal for Philosophy of Science*, 2:119–135.

Infante, P. F., Wagoner, J. K., Rinsky, R. A., and Young, R. J. (1977). Leukaemia in benzene workers. *The Lancet*, 310(8028):76–78.

Kohler, R. (1994). *Lords of the fly: Drosophila genetics and the experimental life*. University of Chicago Press, Chicago, IL.

LaFollette, H., and Shanks, N. (1996). *Brute science: Dilemmas of animal experimentation*. Routledge, London.

Levy, A., and Currie, A. (2015). Model organisms are not (theoretical) models. *British Journal for the Philosophy of Science*, 66:327–348.

Machamer, P., Darden, L., and Craver, C. (2000). Thinking about mechanisms. *Philosophy of Science*, 67:1–25.

Mäki, U. (2001). Models. In *International encyclopedia of the social and behavioural sciences*, volume 15, pages 9931–9937. Elsevier, Oxford, UK.

Nagl, S., Williams, M., and Williamson, J. (2008). Objective Bayesian nets for systems modelling and prognosis in breast cancer. In Holmes, D., and Jain, L., editors, *Innovations in Bayesian networks: Theory and applications*, pages 131–167. Springer, Berlin.

Neapolitan, R. E. (2004). *Learning Bayesian networks*. Pearson / Prentice Hall, Upper Saddle River, NJ.

Pearl, J. (2000). *Causality: Models, reasoning, and inference*. Cambridge University Press, Cambridge.

Reichenbach, H. (1956). *The direction of time*. University of California Press, Berkeley and Los Angeles, 1971 edition.

Rinsky, R., Young, R., and Smith, A. (1981). Leukemia in benzene workers. *American Journal of Industrial Medicine*, 2(3):217–245.

Russo, F., and Williamson, J. (2007). Interpreting causality in the health sciences. *International Studies in the Philosophy of Science*, 21(2):157–170.

Russo, F., and Williamson, J. (2011). Generic versus single-case causality: The case of autopsy. *European Journal for Philosophy of Science*, 1(1):47–69.

Russo, F., and Williamson, J. (2012). Envirogenomarkers: The interplay between mechanisms and difference making in establishing causal claims. *Medicine Studies*, 3:249–262.

Salmon, W. C. (1998). *Causality and explanation*. Oxford University Press, Oxford.

Schulte, P., and Perera, F., editors. (1993). *Molecular epidemiology: Principles and practices*. Academic Press, Cambridge, MA.

Spirtes, P., Glymour, C., and Scheines, R. (1993). *Causation, prediction, and search*. MIT Press, Cambridge MA, second 2000 edition.

Taylor, R. C., Cullen, S. P., and Martin, S. J. (2008). Apoptosis: Controlled demolition at the cellular level. *Nature Reviews Molecular Cell Biology*, 9(3):231–241.

Vineis, P., and Perera, F. (2007). Molecular epidemiology and biomarkers in etiologic cancer research. *Cancer Epidemiology, Biomarkers and Prevention*, 16:1954–1965.

Weber, M. (2005). *Philosophy of experimental biology*. Cambridge University Press, Cambridge, UK.

Williamson, J. (2005). Objective Bayesian nets. In Artemov, S., Barringer, H., d'Avila Garcez, A. S., Lamb, L. C., and Woods, J., editors, *We will show them! Essays in honour of Dov Gabbay*. volume 2, pages 713–730. College Publications, London.

Williamson, J. (2013). How can causal explanations explain? *Erkenntnis*, 78:257–275.

Zhang, L., Rothman, N., Li, G., Guo, W., Yang, W., Hubbard, A. E., Hayes, R. B., Yin, S., Lu, W., and Smith, M. T. (2007). Aberrations in chromosomes associated with lymphoma and therapy-related leukemia in benzene-exposed workers. *Environmental and Molecular Mutagenesis*, 48(6):467–474.

Further Reading

A good introduction to scientific models more generally is given by Frigg, R., and Hartmann, S. (2005). Scientific models. In Sarkar, S., editor, *The philosophy of science: An encyclopedia*, pages 740–749. Taylor & Francis, New York.

More detail on Bayesian networks and causal models is given by Pearl, J. (2000). *Causality: Models, reasoning, and inference*. Cambridge University Press, Cambridge.

Russo and Williamson put forward the epistemological thesis that establishing a causal claim in medicine typically requires establishing both an appropriate correlation and an underlying mechanism. Russo, F., and Williamson, J. (2007). Interpreting causality in the health sciences. *International Studies in the Philosophy of Science*, 21(2):157–170.

Some more information on this thesis is given by the following: Clarke, B. (2011). *Causality in medicine with particular reference to the viral causation of cancers*. PhD thesis, Department of Science and Technology Studies, University College London, London; Clarke, B., Gillies, D., Illari, P., Russo, F., and Williamson, J. (2014a). Mechanisms and the evidence hierarchy. *Topoi*, 33(2):339–360; Gillies, D. A. (2011). The Russo-Williamson thesis and the question of whether smoking causes heart disease. In Illari, P. M., Russo, F., and Williamson, J., editors, *Causality in the sciences*, pages 110–125. Oxford University Press, Oxford; Russo, F., and Williamson, J. (2011). Generic versus single-case causality: The case of autopsy. *European Journal for Philosophy of Science*, 1(1):47–69; Russo, F., and Williamson, J. (2012). Envirogenomarkers: The interplay between mechanisms and difference making in establishing causal claims. *Medicine Studies*, 3:249–262.