Teaching the Tools of Pharmaceutical Care Decision-Analysis

Brian E. Rittenhouse

School of Pharmacy, University of North Carolina, Chapel Hill NC 27599-7360

INTRODUCTION

While there are numerous definitions of pharmaceutical care, (1-4) they all have in common an effort to assist patients and the health care delivery system in optimizing medication delivery and use. As pharmacists become more involved in the delivery of pharmaceutical care, it is vital that they develop new abilities to aid in fulfilling the increased demands that accompany the greater responsibilities undertaken by (or imposed upon) the profession. Historically, medical care decision making has concerned itself with medical outcomes. This is certainly likely to be the emphasis in traditional practices of pharmacy. In the increasingly nontraditional settings in which many pharmacists find themselves, e.g., formulary committees, traditional concerns are beginning to incorporate supplemental information concerning the economics of treatments. Whether they are directly involved in decision making about health care delivery or not, to operate in this new environment, pharmacists will need a greater understanding of how and why such decisions are made.

Regardless of the basis for these decisions (soley health or supplemented by economics), a near pervasive aspect of any medical decision making is the existence of some degree of uncertainty. Many treatments have effects, both health and economic, that are not uniform across the entire patient population. The fields of epidemiology and economics attempt to reduce and quantify this uncertainty. Decision making in the face of such uncertainty is obviously more difficult than in its absence.

In the engineering and business/economics fields (and increasingly in health care), techniques of decision analysis have been brought to bear in order to systematically analyze choices and assist in identifying superior alternatives when decisions must be made under conditions of uncertainty. In health care, these techniques typically meld health and economic outcomes by appealing to epidemiologic and economic evidence as inputs into decision analytic models of treatment choices.

The popularity and sometimes apparent simplicity of these techniques belies their potential complexity. While these methods do assist analysts by simplifying decision making, they cannot work magic. Furthermore, the techniques have a history of sophisticated development that is not always apparent. Adapting the techniques on an ad hoc basis in a seemingly reasonable way may be wholly inappropriate. It is vital that practitioners be aware of the uses and limitations of decision analysis and the boundaries on its appropriate use. This is not always the case. The techniques must be applied intelligently and carefully.

An 18th century social commentator and poet, Alexander Pope, once wrote, “A little learning is a dang’rous thing”. He could have been writing about decision analysis. The under/supply of formally trained decision analysts in the current health care analytical environment has contributed to the self-taught nature of many practitioners in the field. Unfortunately, many of the self-taught (and perhaps others) are not as acquainted with the relevant and essential literature as they need to be. Schechter(5) has indicated several examples of decision analyses in the published pharmacy health care literature which appear to have little acquaintance with proper established analytical procedures. These are not mere theoretical quibbles with methodology; in spirit they are akin to implementing arithmetic addition “theory” by assuming that two plus two equals five. Violations of the theory are simply going to give the wrong answer. The problem is that few readers (or, apparently, reviewers) are sufficiently acquainted with the methods to identify errors prior to publication. The concept of peer review takes on a new connotation in these cases.

It is in part because of this state of affairs that the University of North Carolina at Chapel Hill program in Pharmacy Administration has emphasized formal training of its students by professors trained in specific disciplines like epidemiology and economics who, because of their advanced training in these disciplines, are more aware of the pitfalls that await the uninformed practitioner. While expertise is no guarantor of perfection, some of the more egregious errors of the past could have been avoided by such formal training.

This article indicates the power of decision analysis in two aspects of health care decision making. It shows how epidemiology and economics may be integrated in illustrating standard epidemiologic principles with the techniques of decision analysis. It also shows how treatment choice can be informed and optimized with the appropriate melding of epidemiologic and economic evidence through such “techniques. In the right hands, decision analysis is an illuminating tool that can assist all health care providers and decision makers in optimizing health care delivery. In the wrong hands, it can provide a technocratic illusion of scientific decision making that may violate the axiom—“first do no harm.”

A DIAGNOSTICS EXAMPLE

The following example provides an illustration of both the deceptive nature of medical decision making under conditions of uncertainty as well as the power of decision analysis in assisting the process. It is taken from an article by Eddy(6). An asymptomatic woman goes to her doctor for a mammogram. For purposes of the example, we assume that

the chances of her having breast cancer are one out of 100 (0.01). We also assume that if she has cancer, the test will indicate cancer 79.2 percent of the time, and if she does not have cancer, the test will indicate the absence of cancer 90.4 percent of the time. The question of interest is what are the chances that she has cancer if she has a positive test?

This is worth taking a moment to consider because it is easy to make a mistake. Eddy’s informal survey of physicians indicated that about 95 percent of the physicians estimated this probability at 75 percent. In fact, the correct answer is an order of magnitude smaller (7.8 percent). What is it that explains such an error rate? Eddy’s explanation is that people in general (not just physicians) find conditional probabilities very confusing. Specifically, they confuse the probability of cancer conditional on having a positive test with the probability of a positive test conditional on having cancer. Both are conditional probabilities; they are unlikely ever to be equal.

The problem illustrated in this example is that of casual, unsystematic thinking and, perhaps, a misunderstanding of probability theory. Techniques of decision analysis provide helpful aids in such situations. Decision analysis offers an organized, systematic and explicit diagrammatic technique for organizing thoughts and making computations. All of the issues are laid out explicitly in a diagram where the logical relationships between events and their probabilities can be easily seen. Calculations follow rather easily from the diagram. Decision analysis can be very useful in teaching systematic way of interpreting conditional probabilities and in learning specifically how to answer the question posed earlier to which our group of physicians performed so poorly. The following exposition will take care to build the concepts step by step. At the end of the example, looking back at the basic decision analysis should be much simpler than the length of exposition might indicate. However, the incremental approach is adopted for didactic purposes.

GENERAL PROBABILITIES OF INTEREST

The building blocks of decision trees are probabilities of various kinds. These are described in Figure 1. The area of circle A represents the probability of event A (e.g., having breast cancer in a particular risk group—not drawn to scale for our numerical example). The area in the entire box can be “normalized to equal 1.0” (meaning, call it one and then the make area A the correspond to the appropriate fraction of one equaling the probability of breast cancer). The area in the box that is not part of circle A represents the probability of A not occurring (e.g., not having breast cancer). This probability must equal (1−A) in area (one either has or does not have breast cancer, so the two events exhaust all possibilities).

The area of circle B represents another event probability (e.g., having a positive test with a mammogram). Clearly there is no reason for areas A and B to be equal to each other. The intersecting area (A & B) represents the joint event of both A and B occurring (i.e., having cancer and having a positive mammogram result).

In discussing the concept of conditional probability, we may envision these probabilities as combinations of these areas, A, B and (A & B). In constructing conditional probabilities, one restricts the set of events of interest more than we have done thus far. Let’s say we are interested in the probability of experiencing event B conditional on event A having happened. This is not the joint probability of both events actually occurring, but the probability of B happening if A has happened. Thus it is a proportion not of the entire box, but rather a proportion of circle A. This means that we are to restrict our attention to cases where A has happened (e.g., breast cancer is, in fact, present). The conditional probability of B given (or conditional on) A is a proportion of those with breast cancer (A), just as the original probability (A) is a proportion of all persons in a particular risk group (the box). The conditional probability is not all cases of breast cancer, but only those cases that also test positive [area (A & B)] as a fraction of the number with breast cancer (A).

Note that we could have done all of this restricting our attention to area B instead of A. That is, we could have calculated a different conditional probability (A conditional on B). In this calculation, we would still be interested in the area (A & B), but this time as a fraction of the area B, representing the new different conditioning event. This time the conditional probability of interest is the probability of having breast cancer conditional on having a positive test. The area (A & B) is common to both calculations, but the proportion of interest is different. In one case we are interested in (A & B) as a proportion of A events, and in the other we are interested in (A & B) as a proportion of B events. Since A and B are unlikely to be equal, the conditional probabilities are unlikely to be equal.

SPECIFIC PROBABILITIES OF INTEREST

The prevalence of breast cancer for women of this risk group is 0.01 (meaning one out of 100 women like the one in our example will actually have breast cancer). This prevalence may also be thought of as the probability that the woman has breast cancer prior to any diagnostic event—a “prior” probability. We denote this probability by P(C) for probability of cancer; it is a probability like the area A from Figure 1. The accuracy of the diagnostic test is assumed to be described by a sensitivity of 0.792 and a specificity of 0.904. Sensitivity and specificity refer to two distinct concepts—two “conditional probabilities.”

Sensitivity is the probability that a positive test, indicating breast cancer, will be the result when the test is given to women with breast cancer. Here it is assumed to be 0.792, meaning that 79.2 percent of women with breast cancer will test positive with this test. This probability can be denoted P(+|C), the probability of a positive test given (or conditional on) the presence of cancer.
Table I. Probability notation.

<table>
<thead>
<tr>
<th>Notation</th>
<th>Meaning</th>
<th>Numerical value</th>
</tr>
</thead>
<tbody>
<tr>
<td>P(C)</td>
<td>Probability of cancer (prevalence; prior)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>P(NC)</td>
<td>Probability of no cancer [= 1 - P(Q)]</td>
<td>(0.99)</td>
</tr>
<tr>
<td>P(+/C)</td>
<td>Sensitivity ( = probability of a positive test given cancer is present)</td>
<td>(0.792)</td>
</tr>
<tr>
<td>P(-/C)</td>
<td>(1 - sensitivity) = probability of negative test given that cancer is present = false negative</td>
<td>(0.208)</td>
</tr>
<tr>
<td>P(-/NC)</td>
<td>Specificity ( = probability of negative test given no cancer is present)</td>
<td>(0.904)</td>
</tr>
<tr>
<td>P(+/NC)</td>
<td>(1 - specificity) = probability of positive test given cancer is not present = false positive</td>
<td>(0.096)</td>
</tr>
<tr>
<td>P(+,NC)</td>
<td>Joint probability of having no cancer and a positive test [=P(NC,+)]</td>
<td>(0.095)</td>
</tr>
<tr>
<td>P(+,C)</td>
<td>Joint probability of having both cancer and a positive test [=P(C,+)]</td>
<td>(0.008)</td>
</tr>
<tr>
<td>P(+)</td>
<td>Probability of a positive test result = P(C,+)+P(NC,+))</td>
<td>(0.103)</td>
</tr>
<tr>
<td>P(C1+)</td>
<td>= Posterior probability (probability of cancer given a positive test)</td>
<td>(0.078)</td>
</tr>
</tbody>
</table>

Specificity is the probability that a negative test, indicating no cancer, will be the result when the test is given to women without breast cancer. Here it is assumed to be 0.904, meaning that 90.4 percent of the women without breast cancer will test negatively. This probability can be denoted P(-/NC), the probability of a negative test result given (or conditional on) no cancer.

Both sensitivity and specificity are probabilities of the test giving correct results, in the former case when disease is present and in the latter case when it is not present. We can see that there is no unidimensional concept of “accuracy” in diagnostic testing. Rather, accuracy is two-dimensional. Table I describes the notation used in the example to indicate results from the test along with the interrelationships among these probabilities.

The decision “tree” (the basic diagrammatic construct of decision analysis) describing these events is indicated in Figure 2. The discussion will begin with reference to the far left of the figure (for the moment, ignore the part of the figure to the right of the vertical dashed line). The woman in the example either has or does not have cancer; we indicate this in Figure 2 along with the probabilities of these events equal to the prevalence, P(C), (or prior) and its complement [1 - P(C)]. The circle connecting the two branches of the decision tree indicates the presence of uncertainty; all branches (in general there can be more than two) at such a so-called “chance node” must contain a set of probabilities that sums to equal 1.0. We are given a numerical value (0.01) for the prevalence (and can calculate its complement -0.99).

The decision tree gets “bushier” after additions of the events relating to test results. One can test positive or negative. The probability of testing positive if one has cancer is different from the probability of testing positive if one does not have cancer. These are conditional probabilities, conditional on the events, cancer and no cancer (notationally C and NC). As noted above, the conditional probabilities of particular interest are those represented by sensitivity and specificity. Since each chance node in this example has only two branches coming from it and the probability of one of each is known (sensitivity or specificity), the other two probabilities are also known (since all probabilities at any chance node must sum to equal 1.0). Thus the numerical values for sensitivity and specificity and their complements could be substituted for the notation indicating probabilities in Figure 2. All six of the probabilities from the tree discussed thus far are known. They are calculable from the three initial pieces of information shaded in Figure 1, prevalence, sensitivity and specificity.

At the ends of the branches of the left side of Figure 2 are joint probabilities. These probabilities can be defined notationally by reference to the definition of conditional probability below:

\[ P (+|C) = \frac{P(+/C)}{P(C)} \]

Eq.1

The conditional probability above is equal to a specific joint probability divided by the prior probability of cancer. The probability of a positive test conditional on cancer must be restricted to people with cancer. There are two “ways” to have cancer. One can have cancer and have a negative test [P(C, -)] or one can have cancer and have a positive test [P(C, +)]. The sum of these two probabilities is equal to the probability of cancer.

\[ P(C) = P(C,+) + P(C,-) \]

That is, the restricted universe of persons we want to consider. Then we want to consider how many of them have a positive test [P(C, +)]. Thus, we can see that Equation 1 indicates the proportion of cancers that are positive tests [P(+|C)].

In Equation 1, one may solve for the joint probability in terms of the conditional and prior probabilities: (P(+,C) = P(+|C) P(C). Note that we could have made this point with another conditional probability P(C/+|) as follows:

\[ P(+/C) = \frac{P(C, +)}{P(C)} \]

Eq.2

The above joint probability is equal to P(C/+|)P(+). For joint probabilities the order of the events does not matter; therefore, P(C,+ +) = P(+,C).

Figure 2 shows the additional element of joint probabilities on the decision tree, also indicating the equivalence of P(C,+) to P(+,C). At the moment, it is not clear why this is
helpful, but this will become apparent below. These joint probabilities can be calculated by rearranging Equation 1 to solve for \( P(C,+) \) once the conditional and marginal probabilities are known. In this way, all the joint probabilities in Figure 2 could be calculated. Thus, for example, \( P(+,C) = P(C) P(+|C) = (0.01) \times (0.792) - 0.008 \). Table I indicates the numerical values for all probabilities in Figure 2.

Our ultimate objective is to find \( P(C|+) \) which is nowhere in our figure thus far. We will get this number by a technique called “tree flipping”(7). This involves building a new tree, using elements from the old one, in new combinations. Note that \( P(C|+) \) is a conditional probability [but will not equal \( P(+|C) \)]. We now turn our attention to the right side of Figure 2, specifically the far right.

Here we are looking for probabilities of test results first which are then followed by probabilities of cancer conditional on the test results, the probabilities of ultimate interest for our example. We use the same data but reorganize it. There are only two ways to get a positive test—false positives and true positives symbolized by the joint probabilities \( P(NC,+|) \) and \( P(C,+|) \), respectively. The sum of these probabilities is the probability of a positive test (0.103). Note that we could do the same to obtain the probabilities of a negative test or, recalling that at a chance node all probabilities must sum to equal 1.0, we can simply subtract the above answer from 1.0 and obtain the same result. (This serves as a useful check.)

We now have the basis for a new decision tree and can expand it further into conditional probability branches of cancer (and no cancer) conditional on these test results. These conditional probabilities are unknown (they are what we are seeking in the exercise). In the “flipped tree” we can fill in the ends of the new tree (joint probabilities) with the ends of the old one (since \( P(C|+) \) is equal to \( P(+|C) \), for example (see arrows). Note that the order of the middle two joint probabilities transferred to the new tree will change because of the difference in constructing the new tree causing \( P(-,C) \) and \( P(+,NC) \) to appear in different positions on the new tree.

At this point, everything in the right side “flipped tree” is known except the conditional probabilities of interest. We can now calculate them using the definition of conditional probabilities. Recall Equation 2. The elements on the right hand side of the equation are known and can be used to calculate \( P(C|+) \times 0.078 \).

While the preceding explanation is somewhat long-winded with numerous steps, if the reader will examine Figure 2, it should (hopefully) be clear that the diagram allows one to simplify the discussion remarkably. Recall that the only information available initially were the three probabilities within shaded boxes corresponding to prevalence, sensitivity and specificity. This information using the basic laws of probability can be used to transfer this data to a “flipped tree” that contains the answers to our questions about the probability of cancer conditional on positive or negative test results.

Thus, we have seen that decision analysis can illuminate a basic concept in epidemiology—the interpretation of diagnostic testing. The errors described in the survey done by Eddy can be avoided by using this systematic approach. While there are other ways of describing these concepts, this method has appealed to some students, particularly those less mathematically inclined, primarily because the logic flows rather easily. The decision tree provides a systematic, logical and concise representation of all the data necessary to analyze the problem. Knowing some basic concepts, it would be difficult to err in the calculation. The same cannot be said of non-decision tree based calculations.

TREATMENT CHOICE

Figures 3 through 5 show decision trees in action for a more pharmaceutical-related issue, that of the economic evaluation of treatment alternatives. HA-1A was a product designed to be effective in sepsis (specifically gram-negative bacteremia (GNB) cases). It generally must be administered prior to diagnostic confirmation of GNB. In practice, this means that many patients may receive the product who cannot benefit (they do not have bacteremia or do not have gram negative bacteremia). Decision analysis may be used to indicate the potential problems in making inferences beyond the trial population where there were strict inclusion/exclusion criteria.

Figures 3 shows the trial results in the GNB subgroup (post-administration identified). One can see the survival of the HA-1A plus conventional antibiotics vs. the group with placebo and conventional antibiotics (9). The differential survival was 19 percent (and statistically significant). Unfortunately, while this may be satisfactory to show that FCA-IA does work in a sub-group, the product must be given to a larger group. At best, this will dilute the economic attractiveness of the product if the non-GNB cases with HA-1A react similarly to placebo (more costs, with a diluted, but still
real subgroup effect). At worst, the product’s safety may be called into question (if the placebo group has a better survival picture in non-GNB).

In Figure 4, the entire sample results are reported (including non-GNB). There is still a positive differential in favor of HA-1 A, but it is small. Moreover, it is not clear that mere “dilution” of the sub-group positive effect is all that is occurring. Rather it looks as if there is a borderline negative effect of HA-1 A on non-GNB patients. Figure 5 combines the data in a more extensive format, indicating how many in each treatment arm were GNB. Note that the random allocation of the patients in the trial did not achieve an equal distribution of GNB in each arm (0.40 for FIA-1A and 0.33 for placebo). Both proportions are higher than the expected 0.30 in the general (non-trial) population. Since HA-1 A only works in GNB, inflating the number of GNB patients in the trial would be a good idea if the purpose is to show efficacy in a reasonable sample size. However, such results cannot accurately predict economic consequences unless the same stringency for inclusion/exclusion criteria were to be applied in practice. As a consequence, the generalizability of any economic analysis based solely on the trial results may be questionable.

The overall survival rate conditional on HA-1A was 0.61, but it was based on a weighted average of the conditional survival rates with GNB and with non-GNB patients as the equation below shows. This corresponds to the survival rates in the upper half of Figure 5.

\[
P(S| HA-1A) = P(S| GNB) P(GNB) + P(S| nGNB) P(nGNB) = (0.70)(0.40) + (0.554)(0.6) = 0.61
\]

The number, 0.61, is inflated because of the inclusion of more GNB patients than would be expected in real world applications. Thus, in the real world, fewer than 40 percent of patients will be GNB and (necessarily) more than 60 percent will be nGNB. These two changes will cause the 0.61 number to fall in the equation above as the larger survival rate of the two will now be multiplied by a smaller probability of being GNB (and the smaller survival rate will be multiplied by a larger probability of nGNB).

In the placebo case, we also observed a slightly greater than expected rate of GNB (0.33 vs. the expected 0.30). The adjustment of this weighted average will also have interesting effects on the probability of survival for placebo. However, while this adjustment for real world diagnostics decreased the probability of survival for HA-1 A, in the case of placebo, it will increase it. This can be seen in the equation below which is analogous to the previous equation except that it applies to placebo:

\[
P(S| placebo) = P(S| GNB) P(GNB) + P(S| nGNB) P(nGNB) = (0.51)(0.33) + (0.598)(0.67) = 0.57
\]

Here decreasing P(GNB) to its real world estimate of 0.30 (and increasing P(nGNB) correspondingly) implies the opposite effect on overall survival. This is because the probability that is decreased in this case is multiplied by the lower of the two survival rates (in the weighted average). Correspondingly, the probability that is increased [P(nGNB)] is multiplied by the higher of the two survival rates. This implies that the adjustment will increase the weighted average, raising P(S| placebo) from the observed value of 0.57. These adjustments will cause the differential observed in the trial in overall survival rates (0.61 - 0.57) to shrink. It is worth exploring two extremes of diagnostic abilities. The first is perfect diagnostic so that only GNB patients receive the product. Then the weighted average is simply the GNB result and the differential is the sub-group, 19 percent. In a very rudimentary economic analysis of primary drug treatment costs alone, this would imply an incremental cost of HA-1 A of $21,052 per life saved (based on $4,000 per treatment).

\[
$400,000 \text{ (for 100 patients)} = $21,052
\]

19 lives saved (per 100)

Under a more pessimistic (though not most pessimistic) assumption, if the P(GNB) fell to 20 percent, the overall survival probability differential would fall to zero and the $400,000 would be spent with no benefit resulting in an undesirable infinite incremental cost-effectiveness ratio. This more pessimistic assumption would actually require the percentage of GNB patients to fall below the expected percentage of 30 percent. Poses, et al. have indicated that such possibilities may be very real. In such a case, many more patients receive the product than can benefit and the economics is rather obviously quite unattractive.

**CONCLUSIONS**

The techniques of decision analysis are relatively simple and powerful tools for assisting decision makers in coping with decisions under conditions of uncertainty. The apparent simplicity should not encourage ad hoc experimenting with the methods which can easily violate the axioms of the theory. However, as this article has indicated, these techniques can be of significant value in both illustrating basic concepts of epidemiology and in analyzing more complex treatment decisions and how these can be affected by differential evidence. The power of the tools is tremendous. Misapplication does not diminish the power, just its ability to do good. The training of health care professionals can benefit greatly by the addition of these techniques to the curricula; they need to be taught by those well-trained themselves.

\[2\text{This did, in fact, appear to be the case and largely explains why HA-1 A is not on the market today. (10).}\]
References