Reference: *Deciding Whether to Test Student Athletes for Drug Use*, Charles D. Feinstein, Interfaces, Vol. 20, No. 3 (May - Jun., 1990), pp. 80-87

The role of analytics in decision making is critical when uncertainty is an important aspect of a particular problem. In electric and gas utilities this is especially true because of the uncertain nature of the condition of delivery assets, the influence of asset condition on the likelihood of asset failure, and the resulting impact of such failures on safety and reliability. Diagnostic testing should play and important role in the decision process.

Unfortunately, the state of practice of incorporating test results into the decision process is, at best, poor. The above referenced paper illustrates (1) that analytically incorporating tests into the decision process changes decisions and (2) provides guidance on key parts of the process.

Key aspects of the problem addressed by Feinstein are (1) uncertainty of drug use of individual athletes and (2) the fact that all tests are not perfect. The paper illustrates how to incorporate these two dimensions of the problem, as well as the costs of testing and the costs of decision outcomes, into a formal decision process. This framework could be adopted for use in the electric and gas utility industry. For further discussion of the problem see the links to ***Diagnostic Testing*** and ***Equipment Testing and Decision Models*** in https://www.s-chapel.com.

Feinstein's paper follows:

Deciding Whether to Test Student Athletes for Drug Use

CHARLES D. FEINSTEIN

Department of Decision and Information Sciences
Leavey School of Business
Santa Clara University
Santa Clara, California 95053

The athletic governing board of Santa Clara University had to decide whether to recommend implementing a drug-testing program for our intercollegiate athletes. I presented a decision-analytic model of the question to the board, which served as the focus of the board's deliberations. The deliberations were concerned with evaluating the parameters of the model: the prior probability that an athlete uses drugs, the reliability of the tests for drug usage, and the relative costs of testing errors. These deliberations and the model were the basis of the board's decision not to recommend implementing the testing program.

At Santa Clara University, some straightforward operations research techniques were applied to the analysis of a proposed drug-testing program. I was a faculty member of my university's athletic board of governance when a proposal came to the board from the director of athletics seeking the board's advice and consent to implement a drug-testing program for the university's intercollegiate athletes. In response to that request, I began a detailed study of the question. I developed no new methodology for this analysis, but rather applied some of the techniques of decision analysis [Bunn 1984; Raiffa 1968]. Similar approaches have been employed to study the problem of screening for the human immunodeficiency virus (HIV) [Cleary et al. 1987].

Presenting a Decision Model of the Drug-Testing Issue

When the athletic board of governance

0091-2102/90/2003/0080$01.25
This paper was refereed.

DECISION ANALYSIS—APPLICATIONS

received the proposal, which argued in favor of a drug-testing program, I believed that an analytic approach should be presented that would serve to focus debate on the question. I took it upon myself, as the only decision scientist on the board, to develop such an approach. The reasons I did so were twofold. First, I am opposed to such testing for many reasons, not all of which need concern us here. Second, the athletic director's proposal could not be characterized as analytic but instead was supported by such notions as the necessity to comply with NCAA directives and the putative preventive and deterrent effects of testing. Regardless of how one weighs such considerations, they are surely incomplete without further analysis.

It was clear to me that an appropriate way to proceed would be to develop a simple decision model that treats the question of whether or not to test a single individual for the presence of drugs. Alternative, more aggregate, approaches would not highlight what I believe to be the most important aspect of this question: the consequences of testing errors. Indeed, it is precisely the imperfect nature of the testing that suggests the modeling approach taken. Further, what I hoped that the model would suggest to the board members is that the two essential issues in question were the reliability of the testing procedure and the benefits of identification of a drug user compared with the costs of errors (false accusations of usage and nonidentification of users, the type I and type II errors). I further hoped that the model would demonstrate that these issues were quantifiable. Finally, the model does not address the question of compliance with NCAA guidelines, an important issue, since it seems inappropriate to associate that aspect of the question with a single individual. The NCAA does not require drug testing for all athletes.

I developed my ideas in a memo that I distributed to the 15 board members prior to the meeting. (Approximately one-third of the board are faculty, one-third are university administrators, and one-third are alumni or other interested individuals from the community.) The heart of the argument was contained in a decision tree model (Figure 1) and a table of posterior probabilities of drug use (Table 1).

Figure 1 depicts the decision to test an individual for drug use. The two main alternatives are "test" or "don't test." Following the lower path of the tree, if the alternative "don't test" were selected, the outcome of that decision would be the lottery, governed by prior probabilities, on

> I took it upon myself, as the only decision scientist on the board, to develop such an approach.

"uses drugs" (the individual is a drug-user) and "does not use drugs" (the individual is not a drug user). The prior probability of the event "uses drugs" is the probability that a randomly selected individual is a drug user, which is the fraction of the population that uses drugs.

However, if the alternative "test" were selected, the test(s) would be administered and the results would be obtained from the appropriate laboratory. These results can be "+," that the individual has tested

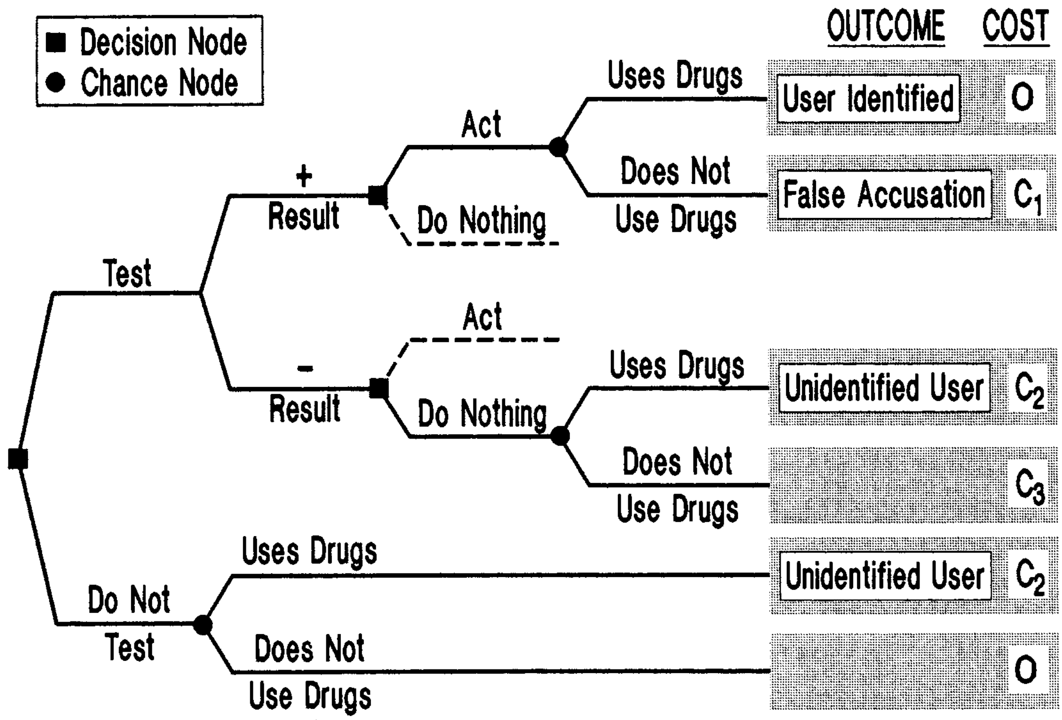

Figure 1: The decision tree model of the question of drug testing evaluates the expected cost of testing for drug use compared with that of not testing. If testing is chosen, the test is given and the result, positive or negative, is observed. If the result is positive, action is taken. Since not all those who test positively are actually users, there is some chance of a false accusation, which costs an amount c_1. If the result is negative, then some drug users are not identified, which costs c_2. Nonusers who test negatively might be expected to experience some cost, c_3, perhaps based on invasion of privacy. The expected cost of the decision to test is computed with respect to the lotteries shown in the tree. The expected cost of not testing is just the cost of an unidentified user, c_2, multiplied by the prior probability that an individual is a drug user. The optimal decision has the lowest expected cost.

positively for the presence of drugs, or "−," that the individual did not test positively for the presence of drugs. This includes the cases of multiple (independent or dependent) tests. The structure of the model doesn't change in such cases, but the likelihoods would vary depending on the kind of testing done. Upon receipt of a positive result (the upper path in Figure 1), the decision maker must now choose to act on this information, "act," or not, "do nothing." The consequences of these alternatives are shown as lotteries involving the true state of drug usage of the individual, "uses drugs" or "does not use drugs." Since the main issue in this problem is whether to test or not, we may as well assume that upon receipt of a positive test report the decision will be to take some action, "act." Hence, in the uppermost path through the tree, I suppress the consequences of the alternative "do nothing" after a positive test result (shown as a dashed line).

In the memo, I then elaborated on the difference between the prior probability

		Prior Probability That Individual Is a Drug User				
		.05	.083	.100	.125	.166
	.99	.84	.90	.92	.93	.99
Reliability	.95	.50	.63	.68	.73	.80
of the Test	.90	.32	.45	.50	.56	.64
	.85	.23	.34	.39	.45	.53
	.75	.14	.21	.25	.30	.38

Table 1: Posterior Probabilities. The probability that an individual who tests positively is actually a drug user depends on the fraction of individuals in the population that are drug users (the prior probability of drug use) and the reliability of the test, which is the probability that the test will be positive if the individual is a drug user. For example, a test that is 95 percent reliable applied to a population that has five percent drug users will yield only a 50 percent posterior probability of drug use. That is, only 50 percent of all those that tested positively will actually be drug users.

that the individual is a drug user and the posterior probability, after observing a positive test result, that the individual is a drug user. The prior probability is just the fraction of drug users in the population. The posterior probability is the fraction of drug users among those who tested positively. That fraction is not one, and is determined by Bayes' theorem, which I introduced in the memo. I made some arguments that specified bounds for the reliability of the test. Tl.e reliability of the test is measured by its sensitivity, the probability that the test is positive if the individual is a drug user. I made the further assumption that the sensitivity of the test is equal to its specificity, the probability that the test will be negative if the individual is not a drug user. This assumption allowed me to create Table 1, a table of posterior probabilities that is easy to read. This assumption, for the range of values shown, is conservative: the assumption tends to overestimate the number of true positives reported by drug tests. Some measurements [Hansen, Caudill, and Boone 1985] have indicated that although the false positive rate, the complement of the specificity, is small, the false negative rate, the complement of the sensitivity, is large. That means that the sensitivity, the true positive rate, is smaller than the specificity, the true negative rate. Hence, the assumption made, setting the two rates equal, overestimates the rate of true positives for reliabilities near one. This, in turn, would tend to overestimate the posterior probability that a positively tested individual is a drug user, which makes the test appear to perform better than it actually would.

Table 1 shows the posterior probabilities as a function of the reliability of the test and the prior probability that an individual is a drug user. To interpret the table, I told the board members that a test that was 95 percent reliable had as much predictive power as the toss of a coin if the population to be tested consisted of approximately five percent drug users. The table was designed to focus attention on the relationship of the population of individuals we were contemplating testing, measured by the prior probability of drug use, the performance of the test, measured by the sensitivity, and the posterior probability that a positively testing individual is actually a drug user.

In the memo I argued that it would be incomplete to judge the worth of the test by probabilities alone. Relying on everyone's intuitive understanding of gambling, I argued that the criterion of expected costs could be used to compare the worth of the two alternatives, to test or not to test. I assigned costs to the six possible outcomes. I assigned the top outcome, the accurate identification of a drug user, zero cost. The second outcome, falsely accusing a nonuser, is the type I error which I assigned cost c_1. The third outcome, doing nothing in the case of a drug user, is the type II error, which I assigned cost c_2. The fourth outcome is not an error but incurs, perhaps, a social cost of this process, which I noted as c_3. This cost might be based on the invasion of privacy a nonuser experiences if compelled to take a drug test. The fifth and sixth outcomes were assigned costs c_2 and 0, respectively. Based on these cost assignments, I developed conditions under which the expected cost of testing was lower than that of not testing and argued that only under such conditions would testing be the rationally preferred alternative.

The simplest necessary condition is that testing is preferred to not testing only if the posterior probability that an individual with a positive test result is a drug user is greater than the ratio of the cost of a false accusation to the sum of the costs of a false accusation and the cost of doing nothing in the case of a drug user, or $c_1/[c_1 + c_2]$. By contraposition, if this condition is not satisfied, then testing cannot be preferred to not testing.

For example, if it were believed that the cost of the Type II error, c_2 (freeing the guilty) and the cost of the Type I error, c_1 (convicting the innocent) were equal, then the posterior probability would have to be at least one-half for testing to be the preferred alternative. If it were believed that the ratio of the costs of the errors were one-tenth, then the posterior probability would have to be larger than 90 percent (actually, 0.9091).

I assigned the top outcome, the accurate identification of a drug user, zero cost.

Clearly, considering the results shown in Table 1, this condition can be used to eliminate immediately some tests under certain conditions. For example, were the ratio of the costs of the errors one-tenth, under none of the situations considered with respect to the prior probability of drug use would a 95-percent-reliable test be acceptable. That is to say, since the cost of the type I error is so large relative to the cost of a type II error, a test that is 95 percent reliable is simply not accurate enough to justify the risk entailed in testing.

I concluded the memo by observing that, based on the model, all that is required to analyze the situation is an understanding of the costs of the various outcomes, which can conveniently be expressed as a ratio. I further noted that in *Zadig*, Voltaire wrote: "It is better to risk saving a guilty person than to condemn an innocent one" [Bartlett 1968]. Sir William Blackstone in his *Commentaries on the Laws of England* (now perhaps regarded more as a literary work than a scholarly guide to jurisprudence), wrote that "It is better than 10 guilty persons escape than one innocent suffer" [Bartlett 1968]. These statements suggest that the cost of the type II error is smaller than the cost of the type I error. One would then conclude, based on the condition, that the posterior probability that an individual is a drug-user given a positive result on the test would have to be larger than one-half for testing to be the rational choice.

The Board's Decision Process

At the meeting called to discuss this issue, the athletic director reviewed his proposal. I reviewed my memo. These were presentations of facts and principles and in no way adversarial. The board seemed very receptive to the model and was appropriately surprised at the results in Table 1. The 50-percent posterior probability for the 95-percent-reliable test was of particular interest. Indeed, the notion of a posterior probability required some discussion, and I had to use several examples to illustrate the concept behind Bayes' theorem. Once I had done so, however, no one questioned the logic and the results. The board discussed what values of the parameters were appropriate for our university. No one suggested that we suffer from a high incidence of drug abuse among our athletes, so that the prior probability that an individual athlete uses drugs was believed to be low, certainly less than 10 percent and perhaps closer to five percent. Most people had read about the reliability of drug tests in various publications and all agreed that Table 1 seemed to capture the range of possible values, with consensus around 95 percent as a representative value. The board further believed that the type I error was indeed more serious than the type II error in this situation, and that closed the case.

The Athletic Board of Governance is an advisory board to the president of the university. Therefore, the board can only recommend policy. In this case, the board voted, with no dissent, to recommend to the president that the university not begin drug testing of our student athletes. The president of the university adopted that recommendation. This analysis was persuasive. The chairman of the board later said to me that the memo was the essential motivation for the recommendation.

Conclusion

I had hoped that my memo would focus discussion on what I believed to be the fundamental analytic issues of the decision to test athletes for drug use. The decision model I presented did just that: the discussion of the board was about what values of prior probability, test likelihoods, and costs of errors were appropriate for our particular situation. Based on this simple operations research model, people analyzed the decision and came to a conclusion that was rational.

APPENDIX

In this appendix, I derive the necessary condition discussed in the text. With respect to Figure 1, define the following notation:

$p\{D\}$ = the prior probability that a randomly selected individual is a drug user;

$p\{D|+\}$ = the (conditional) probability that an individual is a drug user given that the drug test result was positive;

$p\{D|-\}$ = the (conditional) probability that an individual is a drug user given that the drug test result was negative;

$p\{+\}$ = the probability that a drug test will result in a positive outcome;

ACT = the decision to take some action; and

TEST = the decision to test an individual for drug use.

Let ′ denote the negation of a condition; for example, D' is the event that an individual is not a drug user.

Then, the expected cost of the decision *ACT*, conditional upon a positive test report is $E\{ACT|+\} = p\{D|+\}(0) + (1\text{-}p\{D|+\})(c_1)$. Similarly, the expected cost of doing nothing, noted here as *ACT′*, conditional upon a negative test report is $E\{ACT'|-\} = p\{D|-\}(c_2) + (1\text{-}p\{D|-\})(c_3)$.
The expected cost of the decision *TEST* is $E\{TEST\} = p\{+\}E\{ACT|+\} + (1\text{-}p\{+\}) \times E\{ACT'|-\}$.
Similarly, the expected cost of the decision *TEST′* is $E\{TEST'\} = p\{D\}(c_2) + (1\text{-}p\{D\})(0)$ and the decision *TEST* is to be preferred to the decision *TEST′* if and only if $E\{TEST\} < E\{TEST'\}$.
It is straightforward to show that the last inequality is equivalent to

$$p\{D|+\} > c_1/[c_1 + c_2] + (c_3/[c_1 + c_2]) \times [p\{-|D'\}\,p\{D'\}/p\{+\}]. \quad (*)$$

The condition (*) is both necessary and sufficient for preferring *TEST* to *TEST′*. Since the last term in the expression is nonnegative (and may be zero if c_3 is set to zero), it is convenient to derive a necessary condition from (*) by simply dropping the last term. Thus, if *TEST* is preferred to *TEST′*, it is necessary that the following holds:

$$p\{D|+\} > c_1/[c_1 + c_2] = 1/[1 + c_2/c_1]. \quad (**)$$

By contraposition, if condition (**) is not satisfied, then *TEST* cannot be preferred to *TEST′*.
Further, the posterior probability $p\{D|+\}$ is determined, by Bayes' Theorem, from the prior probability of drug use and the likelihoods—the sensitivity, $p\{+|D\}$, and the specificity, $p\{-|D'\}$—of the test

$$p\{D|+\} = p\{+|D\}\,p\{D\}/[p\{+|D\}\,p \times \{D\} + (1\text{-}p\{-|D'\})(1\text{-}p\{D\})].$$

Thus, upon selecting the appropriate cost ratio, condition (**) determines the appropriateness of a test with known sensitivity and specificity applied to a population with incidence $p\{D\}$ of drug use.

References

Bartlett, J. 1968, *Familiar Quotations*, 14th edition, Little, Brown and Co., Boston, Massachusetts.

Bunn, D. W. 1984, *Applied Decision Analysis*, McGraw-Hill, New York.

Cleary, P. D.; Barry, M. J.; Mayer, K. H.; Brandt, A. M.; Gostin, L.; and Fineberg, H. V. 1987, "Compulsory premarital screening for the human immunodeficiency virus," *Journal of the American Medical Association*, Vol. 258, No. 13 (October 2), pp. 1757–1762.

Hansen, H. J.; Caudill, S. P.; and Boone, J. 1985, "Crisis in drug testing: Results of CDC blind study," *Journal of the American Medical Association*, Vol. 253, No. 16 (April 26), pp. 2382–2387.

Raiffa, H. 1968, *Decision Analysis*, Addison-Wesley, Reading, Massachusetts.

James F. Sepe, Associate Professor, Leavey School of Business Administration, Department of Accounting, Santa Clara University, Santa Clara, California 95053, writes: "At that time, May 1987, I was a member of the Board, and I remember the circumstances well.

"The athletic director submitted a proposal for beginning drug testing of our athletes. Not only was this proposal justified in terms of compliance with NCAA guidelines, but there were claimed educational and health benefits as well. Before the meeting at which the board was to make its recommendation to the president of the university, the chairman of the board distributed Professor Feinstein's memorandum on the question.

"It is fair to say that the memorandum became the agenda of the meeting. The decision analytic approach provided the framework that the board needed to debate the costs and benefits of the policy. After a bit of explanation of the mathematical foundation of the model—especially the workings of Bayes' Theorem and the idea of a posterior probability—the board began to discuss what values of the parameters were appropriate for our situation. It should be stated that the structure of the model and the approach taken were never questioned and were accepted by all as a reasonable way to formulate the problem. The board concluded that the incidence of drug use on our campus was so low and the cost of a false accusation so high that drug-testing was not the correct decision for us. What was particularly pleasing about this outcome was that we never had to enter into the potentially contentious area of privacy rights and the legalities of drug testing. The model allowed us to resolve the issue clearly, using an approach that could be tailored to our particular situation. There was no dissent in the recommendation."