Acceptance of Actuarial Decision Aids in Medicine

Hila Calev

Princeton University

April 19, 2010

Honor Code:

This paper represents my own work in accordance with University regulations.

Hila Calev

Abstract

Many actuarial decision aids have been developed in medicine that have been shown to improve medical decision-making and patient care. Nevertheless, decision aids are underutilized in medicine, and this reflects a general discomfort that people have with utilizing a formula in place of human judgment. Psychological research has attempted to understand this general reluctance to employ a decision aid, and findings indicate that overconfidence of the physician, as well as fear of derogation by the patient, lead physicians to avoid utilizing actuarial decision aids. When viewed through the framework of Prospect Theory, the reluctance can be attributed to risk-aversion. This is because physicians maintain a reference point in which patients are treated properly and accurately. Actuarial decision aids, which have explicit limitations on their success rate, are therefore seen as an unfavorable sure loss.

Between the years of 1949-1951 and 2003, the average life expectancy in the United States rose from 68.1 years to 77.5 years, an increase of almost 10 years (Shrestha, 2006). Undoubtedly, medical and other scientific research contributed to this trend and both have come very far in improving our health. Going to the doctor is a routine part of our existence, and most of us trust in medicine to make us better. The past few decades have seen a proliferation of medical and scientific research that has fundamentally changed the ways we can do surgery, augmented the number and variety of drug treatments available, and much more. This depiction of the medical practice, however, neglects a fact which psychologists and advocates of evidence-based practice have focused on in recent years: the practice of medicine still involves a significant amount of uncertainty. Diagnoses, including the results of pathological and radiological tests, and prognoses are rarely certain. There are false positives and false negatives. Furthermore, like all human beings, medical decision-makers are subject to various cognitive heuristics and deviations from normative behavior that lead to reduced consistency in their judgments. For example, a physician given the same patient file at two separate times might reach different conclusions based on the same data. These random fluctuations result in sub-optimal decision-making and accuracy (Dawes, Faust, & Meehl 1989).

How would you feel if you walked into your doctor's office, but instead of telling you right away what his opinion was of your condition, he opened up a computer program and advised you according to what it said? How would you feel if you were the doctor, and you did this in front of your patient? Actuarial decision-making aids, like a computer programs, are based on *data;* they are developed through statistical analysis of cases with known outcomes to distinguish diagnostic form non-diagnostic information to weight different diagnostic information appropriately (Swets, Dawes & Monahan, 2000). Actuarial decision-making systems

often outperform even the experts because they are immune from the fluctuations in judgment that human reasoning is subject to; given the same data, an actuarial formula will always produce the same response.

The following paper concerns itself with actuarial decision-making in medicine. To aid physicians and other medical decision-makers in making choices, decision support systems have been developed for clinical treatment and diagnosis. They exist for specific scenarios, such as whether radiography should be done on patients with acute knee or ankle injuries (Graham et al., 2001) or more general cases, like diagnosis of emergency room patients with abdominal pain (De Dombal, Leaper, Horrocks, Staniland, & McCann, 1974). Reviews of a large number of clinical decision support systems report that the majority improves the performance of the medical practitioner. In 1998, a review of 66 studies on clinical decision support systems found that 43, or 66%, improved the performance of the medical practitioner (Hunt, Haynes, Hanna, & Smith, 1998). A more recent review in 2005 similarly found that 62% of the 97 studies analyzed resulted in improved clinician performance, including 40% of the diagnostic decision support systems (Garg et al. 2005). Finally, in the social sciences, Dawes et al. (1989) reports that in 100 comparative studies of actuarial and clinical methods the actuarial method was almost always equal or superior to the clinician.

And yet, despite the evidence that many decision aids improve accuracy, they are drastically underutilized in medicine (Kaplan, 2000). Even though the stakes in medicine are high, physicians will rely on their own intuitions about a case rather than employ a helpful decision aid - at the expense of increased accuracy and better patient care. In the following sections, I present research findings from psychological studies that attempt to explain this reluctance of physicians to employ helpful decision-aids. The findings include overconfidence on

the part of the physician that leads people not to accept decision aids whose accuracy is less than perfect, and a (well-grounded) fear of derogation by the patient.

In the final section of the paper, these same findings are interpreted within the framework of Prospect Theory, as described by Kahneman and Tversky (1984). The purpose of this discussion is to place the reluctance to utilize decision aids in medicine within a larger scheme of the cognitive strategies that physicians use when dealing with patients. Overconfidence, a preference not to accept accuracy below perfect, and fear of derogation are phenomena that may explain the reluctance to employ decision-aids. However, these phenomena may in turn arise from prospect theory-like effects, when doctors interpret their practice through a gain-loss framework.

Arguments for and Against Actuarial Decision Aids

Actuarial decision aids are often better than clinical judgment alone, but arguments have been advanced both in their favor and against their use. Effective actuarial formulas achieve accuracy that is greater than that achieved by even the most expert physicians. They help to sort and manage large amounts of information, and it has been argued that they are necessary for implementing with greater efficiency protocols developed in recent research. Arguments against actuarial methods include the fact that they may systematically fail under certain circumstances and that they may disrupt patient-doctor interactions. These advantages and disadvantages are discussed below.

Advantages of Actuarial Decision Aids

The foremost argument for the use of actuarial decision aids is that they are more accurate than human judges because they always reach the same conclusions based on the same set of data. Even if human judges understand which data are important to consider when making

a decision, they are susceptible to heuristics and biases that prevent them from having the same consistency (Dawes et al. 1989). The fact that these biases are often controlled by automatic, unconscious processes (processes under System I control) means that physicians are unaware of, and therefore unable to correct for, these types of errors (Evans, 2003).

One revealing example is a study by Redelmeir and Shafir (1995) in which the number of treatment options affected whether physicians recommended surgery to a patient or not. Family practitioners were given a scenario in which a hypothetical patient with osteoarthritis had tried several medicines to no avail. Half of the physicians were told that two options remained: one other medication and surgery. The other half were told that three options remained: the same medication and surgery that the other group was given, plus an additional medication. Results showed that 72% of the physicians referred the patient to orthopedics when there were two medications to try, while only 53% referred the patient to the orthopedics when there was only one option. Even though having yet untried medications meant that there was a greater chance that the patient could avoid surgery, a smaller percentage of doctors recommended medication in the second condition. The number of alternatives changed the treatment plan; in order to avoid deciding between two medications, doctors defaulted to the third (much more drastic) alternative. Other factors, such as the severity of the cases the physician has recently seen, (Shen, Rabinowitz, Gesit, & Shafir, 2010) and framing (Kahneman & Tversky, 1984) also lead to deviation from normative principles of decision-making. Cognitive biases such as these, in which physician judgment is affected by inescapable aspects of the work (previous patients, numerous alternatives) are truly unavoidable, even with the best training. Actuarial formulas, on the other hand, behave as a normative decision-maker would in these scenarios.

Dawes et al. (1989) present several other innate shortcomings of human judgment that may lead to errors in even expert judgment. These include a poor ability to distinguish between valid and non-valid predictors and to determine the associations *between* predictors, the tendency to succumb to self-fulfilling prophecies, and the tendency to seek out only information that confirms one's initial intuition or hypothesis. The errors of cognition described above illustrate nothing other than the fact that physicians are no less immune from cognitive heuristics than the rest of humanity. Decision aids derived from actuarial processes are a method of overcoming innate fallibility of human judgment.

Even in the case that a physician has all the tools to make proper decisions, and even when factors such as previous patients and framing of the situation do not come into play, the amount of medical knowledge that a physician must keep track of is enormous. Another advantage of medical decision aids is that they help to manage information, to prevent mistakes of oversight in treatment or diagnosis (Komaroff, 1982). In fact, the most successful category of clinical decision-support systems studied by Garg et al. (2005) in their large-scale review of 97 studies were those that also provided reminders to healthcare practitioners (76% of these systems were shown to improve patient care). Advocates of evidence-based practice in medicine argue that actuarial decision aids are also advantageous because they can increase the rate at which new practices developed from recent research are implemented. The goals of evidence-based practice are undermined without methods of effectively applying new recommendations for care based on recent research. Spring (2008) argues that decision support systems can be particularly helpful in complex sequential clinical management decisions. These decisions must be made as "systematic, deliberative, and computational" as possible in order to manage large amounts of new information and reduce judgment error (Spring, 2008, p. 4).

Disadvantages of Actuarial Decision Aids

Arguments against actuarial decision aids fall mainly into two categories. The first is that there are select circumstances in which human reasoning may outsmart the judgment of an actuarial formula. The other branch of criticisms is not based on the intrinsic nature of the actuarial system but on the claims by physicians that the decision aids are unnecessary and disruptive, with the inconveniences outweighing the benefits. This second category of criticism is highly problematic because it leads to reluctance to use decision aids even when evidence has shown their efficacy.

In their review in support of actuarial methods, Dawes et al. (1989) also present three situations in which the clinical method may systematically yield better predictions, which raises certain concerns about the efficacy of actuarial methods in these cases. The first is in situations when predictions are based on theories and therefore cannot be made based on statistical frequencies alone. Here an actuarial formula may not suffice. The second case occurs when there is an event so rare that it is not accounted for by the statistical method, such as that caricatured by the "broken leg" example. A statistically-derived formula will fail to predict that a man who reliably goes to a movie every week will not go the week he is immobilized by a broken leg. This suggests that clinicians should be able to override actuarial judgment when this rare type of scenario arises. (However, Dawes et al. [1989] also report that when left to their own judgment, clinicians tend to override the actuarial formula too often, and hence have inferior accuracy than the formula anyway.) Finally, the authors point out that humans may have more complex pattern recognition than computer systems in some cases.

Remaining complaints about decision aids cover a variety of issues. Physicians themselves have made the claim that decision support systems are *demeaning* to physicians by

preventing them from exercising their skills (Komaroff, 1982). Another study found that the use of the decision aid disrupts patient-doctor conversation and is not tailored to the specific patient's case (Wendt, Knaup-Gregori & Winter, 2000). (However, as discussed earlier, physicians who tailor decisions too much to patient-specific information tend to do poorly in determining and weighting diagnostic criteria.)

Overall, research shows that actuarial decision aids can clearly be powerful tools in medicine. However, once an effective decision aid has been developed it must be applied properly and consistently in order to improve accuracy, and many physicians are reluctant to do so. Although decision aids have a few legitimate shortcomings, which have been mentioned above, it appears that many physicians are still uncomfortable with using these aids, despite knowing that they can improve patient care, even including situations in which they are disadvantaged (Bouaud et al., 1998). At first, then, the phenomenon of underutilization seems enigmatic. Psychological research has centered on the general reluctance to use actuarial formulas, and several studies have targeted underutilization specifically in medicine. Researchers have attempted to understand any underlying cognitive mechanisms that prevent reliance on actuarial methods. The following section presents the findings of two prominent branches of research in this area.

Research findings on Reluctance to Utilize Actuarial Decision Aids

To understand the directions which research on utilization of actuarial formulas has taken, we return to the two questions posed initially in the opening paragraphs of this paper. How would you feel if you were a doctor who used a computer program instead of giving your own unaided conclusion about the patient you just examined? How would you feel if you were the patient whose doctor used a computer program to decide how to treat you? Research has

investigated the utilization of actuarial formulas in general, and specifically in the context of medicine. Importantly, utilization of decision aids has been investigated from the perspective of both doctor and patient. The two general phenomena that have been examined are overconfidence by the medical practitioner who should defer to the actuarial formula, and the finding that patients simply regard information from an actuarial source differently and therefore hold clinicians who defer to actuarial decision aids in lower esteem.

Overconfidence

A well-established psychological finding is that people are excessively optimistic in their predictions and that this leads to overconfidence in their own abilities (Armor & Taylor, 2002). Indeed, Sieck and Arkes (2005) found a relationship between subjects' defiance of an actuarial formula and overconfidence in their own accuracy. Participants were provided with information about prospective jurors and asked to predict whether or not the juror was in favor of physician-assisted suicide. The experimenters used real data and derived an actuarial formula that achieved a predictive accuracy of 77%. Participants were informed of the formula's accuracy and for each judgment were able to click a button to see the prediction of the formula before making their own prediction. After making their prediction, they indicated how confident they were in it. It was made clear that in providing a different answer than the formula the subjects believed the formula was wrong while they were correct. Results showed that participants performed significantly worse than the formula and over-predicted their own accuracy in cases when they had defied the formula.

In a subsequent experiment in the same study, strong calibration feedback reduced overconfidence when subjects made estimates with no formula present. Subjects had to compare their confidence and actual performance and describe how much they needed to adjust their

confidence to be well-calibrated. When a formula was provided, subjects relied on it more and performed better. Thus, when overconfidence was reduced, people deferred more to the formula. Although the experimenters themselves never considered the point that participants may have been distrustful of the formula itself, confidence manipulation does provide evidence for the hypothesis that it was overconfidence in their *own* abilities that led participants not to use or heed the formula. Overall, the general finding suggests that when an actuarial formula is available, overconfidence discourages its utilization.

One aspect of this study that is important to note in connecting these findings with the medical field is that overconfidence was difficult to reduce; only intense calibration feedback was effective. Indeed, in the same study, the Sieck and Arkes (2005) report that when they had failed to provide strong enough feedback, overconfidence was not reduced. (Therefore, the authors refer to the "recalcitrance of overconfidence" in the paper title.) Medicine is a field in which any calibration feedback at all is often difficult to obtain. For example, doctors cannot be sure whether patients are adhering to treatment plans, cases often play out over extended periods of time, or patients may switch doctors. Indeed, one early study about actuarial decision aid usage in medicine highlights a relevant phenomenon. De dombal et al. (1974) provided clinicians with feedback from an actuarial formula on the diagnosis of abdominal pain in a real surgical unit for about a year and a half. The feedback was provided at the end of each day, regarding the patients who had been seen during that day. The performance of the clinicians improved over time during the course of the study and the authors attributed this effect to the calibration provided by the actuarial program that made the physicians more systematic. Hence, in medicine, effective decision aids may provide one of the only convenient means of calibration. If

overconfidence prevents physicians from using the decision aids in the first place, then overconfidence is in fact especially hard to overcome.

Finally, Arkes, Dawes and Christensen (1986) published findings illustrating other aspects of overconfidence, although the actual participants in their study were not physicians. In their first experiment, incentives proved detrimental to judgment. Participants were asked to predict whether students had graduated with Honors. They were shown three courses that each student had supposedly taken and the student's grades for those courses. Participants were also given a very simple decision rule that would result in 70% accuracy, which was based simply on the number of As the student had received. Groups that were given incentives to perform better tended to rely less on the decision rule and performed worse. In a second experiment, experts performed worse than non-experts. The task required students to predict whether a baseball player had won the Most Valuable Player award for that year, based on several statistics about players. They were told that if they would be correct 75% of the time by simply relying on only one criterion (which team ranked highest in the standings). Results showed that those participants who had expertise in baseball were less likely to use the decision rule and were less accurate than the rule and than the non-experts. They were also overconfident. The authors conclude that incentives and expert status prompted the participants to abandon the decision rule because they did not want to be limited by its success rate, and they believed that they could outperform it. Physicians notably have both high incentives and expertise, and therefore may be prone to overconfidence.

In this study, why did incentives lead participants to abandon the decision rule? Participants in experiment one received outcome feedback after every student whose Honors status they predicted. In order to consistently apply the actuarial rule, the participants needed to

heed it even though it led to an incorrect prediction 30% of the time. The authors argue that incentives decreased the participants' tolerance of the negative outcome feedback and thus led them to abandon the rule.

Another complication that the authors of this article observe but do not discuss in detail is that whoever decides to follow the rule and tolerate the negative outcome feedback also has to accept its fixed and "blatantly imperfect" success rate (Arkes et al., 1986, p. 163). Participants in the control group, which was one of the groups that least heeded the formula, were told that the 70% success rate of the rule was "about as well as people can do… we haven't found evidence that people really can do much better than the 70 percent" (Arkes et al., 1986, p. 167). If people had considered this success rate a benefit – a guaranteed maximum success or even an improvement over their unaided performance – they should have been incentivized to use the formula. However, had they viewed this formula as *limiting* their potential success then using it would not have been preferable to rely on it. It will be discussed later on that in the latter case, in medicine, physicians maintain a frame of reference of fully successful patient treatment, so that every incorrect answer is viewed as a "loss." In the final section of this paper, I discuss how this reference frame reflects the idea that Prospect Theory-like effects may account for some of the behavior of physicians when they are faced with "blatantly imperfect" decision rules.

Derogation of Physicians who use Decision Aids by Patients

Another hypothesis for the reluctance of physicians to utilize actuarial decision aids is that physicians fear derogation by patients, and several studies have shown that this may be a well-grounded fear. Indeed, trust in actuarial decision aids seem to have opposition from both physicians and patients, with the opposition of the patients further contributing to the opposition of the physicians.

Several studies have shown that people trust actuarial methods less than they trust human experts. Onkal, Goodwin, Thomson, Gonu and Pollock (2009) conducted a study to compare the extent to which participants would utilize advice if they believed it had come from an expert or from a statistical model. The participants showed a bias against the models (or a bias toward human experts). The experiment was carried out in the field of stock market forecasting, in which there is no evidence that so-called "expert opinion" is a valuable predictor of results. Participants were provided with stock information and asked to make a forecast. They were then shown advice, and asked to adjust their forecast. One group was told that the advice came from a statistical model and the other group was told that it came from a human expert. Results indicated that if the advice was allegedly from a statistical model, the initial forecast was adjusted less than if it allegedly came from a human. In a second experiment, two sources of advice were provided, from both a human expert and a statistical model. When these sources differed, participants paid more attention to the advice from the human expert.

In a different study, Promberger and Baron (2006) showed that subjects preferred to follow the recommendation of experts over computers because heeding the expert enabled them to feel less responsible for the decision than if they had simply heeded a computer. Subjects were asked to imagine themselves as patients with a particular set of symptoms and they were given a recommendation of whether or not to have surgery, either from a computer or a physician. Patients were asked whether they would heed or defy the recommendation, and then reported how responsible they felt for the decision. In their second experiment, participants reported trusting the recommendation of the physician more than that of the computer. This again provides support for the hypothesis that patients interpret the advice of computers in a different way than the advice of physicians. Promberger and Baron (2006) hypothesize that people prefer

to heed human experts because other humans are capable of being loci for bearing responsibility, thereby relieving the person of some responsibility for the decision, while a computer cannot fulfill such a function.

Finally, several studies have shown that patients perceive doctors who use decision aids to have lesser abilities than those who do not. Cruickshank (1985) reports this finding in a study of three real doctors who used computer aids while seeing 140 hospital outpatients. When using the computer, patients' ratings of the doctors did not match up with the patients' depictions of the ideal doctor as they did when the doctor did not use the computer. This effect was due mainly to perceived inadequacy in the social interactions between the doctor and patient. In a more recent study (in a time when computers are familiar and accepted in general society) Arkes, Shaffer, and Medow (2007) showed that perceived diagnostic ability of physicians who use decision aids was lower than if they did not use a decision aid. The study was conducted with students who read a scenario in which a physician diagnosed them either after using a diagnostic support system or not, and with patients in a waiting room who saw a physician who did or did not use a diagnostic support system. Participant ratings of the diagnostic ability of the physician showed that physicians who used the decision aid were rated to have lower diagnostic abilities. The authors ascribe this to the participants' partial attribution of the blame or the credit for the diagnosis to the decision aid rather than to the physician. This also held in their fourth experiment, in which participants were third- and fourth-year medical students.

Are doctors aware of the biases that patients have for trusting human experts over decision aids? The fact that the effect in the above study by Arkes, Shaffer and Medow (2007) held when the participants were medical students suggests that this faith in the expert human opinion is present in doctors-in-training. Thus, doctors themselves, even if they hold positive

attitudes toward decision aids, should be aware of patient biases. (They may also share these biases.) Evidence for this claim comes from a study by Pezzo and Pezzo (2006) in which medical students correctly predicted the harsh attitudes of undergraduates in regarding a doctor who had made a mistake using a decision aid. In experiment two of this study, participants were either medical students or undergraduates. They were asked to read a scenario in which a physician had made an incorrect diagnosis resulting in the patient's death, and to then rate how blameworthy the doctor was. The doctor had either used or not used a decision aid. If the physician had used the aid, he had either agreed with it, disagreed but heeded it anyway, or disagreed and defied it. Results showed that in the case that the physician had agreed with or heeded the aid, the participants assigned him reduced fault for the error. Thus the decision aid was protective if the doctor abided by it but not if he had defied it. Overall, medical students were less harsh than undergraduates in assigning blame. The medical students were also asked how they thought the undergraduates would answer. Their responses showed that the medical students could predict with reasonable accuracy the harsher attitudes of undergraduates, and they even overestimated how low the undergraduates would rate the physician's competence. In other words, if medical students' opinions are generalizable to full physicians, the results of this study suggest that doctors are aware of and anticipate patient's negative responses to use of decision aids, even to the extent that they overestimate patients' negative responses. This study and the above study by Arkes et al. (2007) suggest that physicians, like patients, may believe that physicians who use diagnostic aids have reduced diagnostic abilities, while at the same time they imagine their patients bearing this same opinion - but even more strongly than they do.

Arkes, Shaffer and Medow (2008) also conducted an experiment on the opinions of mock jurors on a physician who had made an error with or without using a decision aid. The use of the

aid had no impact on the verdict of the case, but the physicians who had defied the aid were deemed deserving of harsher punishment.

A review of these three studies just discussed shows that the pros and cons of using a decision aid for a physician, in terms of how much he will be derogated, is as follows: A physician who uses a decision aid is given less credit for correct diagnoses. When he is wrong, having heeded the aid results in a less negative opinion by the patients, while having defied the aid may cost him greater liability in a malpractice lawsuit. This implies that in order to avoid blame in the case of committing an error, physicians should use an aid and not defy its recommendations. As discussed in the first section of this paper, in general this would likely improve the physician's overall accuracy to the level of performance of the decision aid!

Prospect Theory and the Gain-Loss Framework

Although overconfidence and fear of derogation by the patient are seemingly unconnected phenomena that were discovered empirically to play a role in physician willingness to utilize actuarial decision rules, they can both be placed within the same wider framework to explain physician behavior. The central claim of this final section is that physicians tend to use a cognitive reference point that leads them to view incorrect diagnoses and improper treatment of patients through a loss frame, as described by Prospect Theory (Kahneman & Tversky, 1984). Loss aversion, which arises from Prospect Theory's value function (and will be explained later in this section), explains or corroborates many of the effects that have been discussed in the existing research. If physicians are considered to operate in a loss mentality, a new understanding of utilization of decision rules is made available. There are two major implications of the gain-loss framework. First, the reluctance to rely on decision rules stems from cognitive mechanisms that are very deeply rooted and therefore the phenomenon is pervasive and ubiquitous. Second,

behavior may be altered by framing, based on whether situations are approached through a gain frame or a loss frame, and this may be useful in changing people's attitudes toward decision aids.

Prospect theory is a descriptive theory based on observations that people do not behave in an entirely normative manner when dealing with risky choices in which the consequences are not known, otherwise known as gambles. Kahneman and Tversky (1984) state that people behave according to an S-shaped value function, in which outcomes are encoded as gains or losses from a reference point instead of as states of wealth. The gain function is concave, while the loss function is convex and steeper than the gain function. The function reveals that people display risk aversion in gains, risk-seeking behavior in losses, and loss aversion when one stands either to gain or lose in a gamble.

It follows from Prospect Theory that people must adopt a reference point, from which all subsequent exchanges are seen as gains or losses. This is important because the same scenario may be encoded as a loss or a gain, depending on the reference frame initially adopted. In the classic example, Kahneman and Tversky (1984) provide the scenario of a new Asian disease. Subjects are told that this disease will kill 600 people if no action is taken. Two courses of action are proposed. In plan A, 200 people are surely saved, and in plan B there is a one-third chance that everybody will be saved and a two-thirds chance that no one will be saved. According to this phrasing, the reference point is 600 deaths, and the plans are presented as "saves" or "gains" from this reference point. Therefore the risk-averse option (plan A) is selected by the majority of respondents. The same options can be phrased alternatively so that even though they imply the same outcome, the options are encoded as losses from a reference point of 600 survivors. In this case, plan C is that 400 people surely die, and plan D is that there is a one-third probability that no one will die and a two-thirds probability that all 600 will die. In this case the reference point

causes one to adopt a loss framework, and a majority of people adopt the risk-seeking plan (plan D).

As has been previous mentioned, reluctance to utilize a decision aid can be explained if physicians view every patient who is not treated properly through a loss frame. In other words, physicians adopt a reference frame in which patients are treated properly, and any deviation from this standard must necessarily be viewed as a loss. However, we have also seen that prestige also influences a physician's decision-making. Prestige may also be viewed through a loss frame. The losses and gains which physicians must monitor are described as follows. It is important to note that, according to this description, doctors do not monitor their practice as a whole; instead, they focus on each case on an individual basis.

- Proper or improper treatment of a patient is viewed by the physician as the "currency" of his trade. Each patient is like an investment, in the sense that the doctor puts in time and attention, with the expected return of being able to successfully discharge the patient. The patients are like a currency, and the doctor has a stake in it. Holding a reference point of proper treatment for each patient implies that patients who are misdiagnosed or mistreated are seen as a "loss." Viewing the treatment of each patient a gain or loss from the pool of current patients may be a useful cognitive strategy for physicians to keep track of their patients.
- Physicians must also be concerned with their prestige in the eyes of others. A physician's prestige affects how much respect a patient has for him or her, and therefore how likely the patient will be to follow the physician's recommendation. Prestigious doctors also attract more patients. Prestige is therefore a highly valuable for physicians to possess, both for the financial success and possibly to elicit cooperation from the patient. Given its

importance, it is also plausible that physicians monitor the prestige afforded them by patients on a patient-by-patient basis, each one seen as providing a gain or loss from the physician's overall prestige.

The most important feature of the actuarial decision rule for this interpretation is that it provides a fixed maximum success rate, of which the physician is aware. The fixed rate of failure of the actuarial method is especially salient, so that the physician encodes any success rate that is less than 100% as a *sure loss,* as in plan A in the Asian disease example. In other words, the physician adopts a reference point in which all patients are diagnosed and treated properly. Treating patients is therefore seen as in a loss frame, and this leads to risk-seeking behavior. Physicians are then faced with the choice of accepting the decision rule's verdict or taking a chance by relying on their own judgments, for which there is no clearly defined limit that makes salient their likelihood to fail[1]. The risk-seeking choice is therefore not to utilize the formula, and this is the phenomenon we observe. An exactly analogous situation scenario occurs in the case of prestige. Overconfidence, which in previous research was shown to be related to a reluctance to use decision rules, contributes to this effect by encouraging the physicians that the risk that they take is not as great as it is in reality. Thus, overconfidence contributes to the effect predicted by Prospect Theory by lessening the perceived risk in the risk-seeking option, thus making it more desirable.

The final piece of Prospect Theory describes how people weight differing probabilities. People tend to overweight low probabilities, underweight high probabilities, and have very unstable opinions near the endpoints, when probabilities are extremely low or extremely high.

[1] Although physicians may have a general sense that they do not achieve perfect accuracy, self-serving biases and overconfidence, as well as scarce outcome feedback make it easier for the physician to believe that their own success is not limited in the very clear way that the formula's success is limited.

Kahneman and Tversky (1984) also found that changes around the endpoints, such as moving from a 0% chance to a 1% chance, and moving from a 99% chance to 100% (certainty) are overweighted compared to changes in the middle of the scale, say from 33% to 34%. Actuarial formulas often offer physicians the opportunity to improve their accuracy by only some modest amount that leaves them in the middle of the scale (say, an improvement from 68% to 73%) whereas *theoretically,* a highly overconfident physician may judge the powers of his own intuitive reasoning to be unbounded, with the possibility of reaching 100% accuracy. This helps to corroborate the findings of Arkes, Dawes and Christensen (1986). The idea of possibly reaching 100% accuracy by abandoning the decision rule is particularly enticing to those who are highly incentivized to perform. Thus, physicians and other decision-makers who operate under conditions in which the stakes are high or the incentives are significant are tempted by the idea of perfect accuracy, because attaining certainty is drastically over-weighted according to Prospect Theory.

Consideration of the physician's position in terms of a gain-loss framework has implications for ways to encourage utilization of decision rules. Since according to Prospect Theory people are sensitive to framing effects that change the reference point so that interactions are seen as either gains or losses, situations can be manipulated to risk-seeking behavior. If indeed physicians behave according to this gain-loss framework, then framing effects may be utilized to change a physician's opinion on the utilization of an actuarial decision rule, by framing the decision not as a loss but as a gain. One simple way of changing the framing before presenting the choice to use an actuarial method to physicians is as follows: "A decision aid has been developed that catches X% of mistakes that clinicians make." This presents an improvement from a reference point that is easier to accept. The actual accuracy of the clinician

is not described, and therefore the challenge to reach 100% accuracy is not emphasized as greatly. The success rate of the actuarial rule is coded as a gain, since it reduces an undesirable, "negative" value (analogous to reducing a debt).

Conclusion

Actuarial decision aids are an important tool in medicine whose underutilization may come as a surprise at first. Medicine seeks to be rigorously based in scientific research, and research has proven the benefits of many actuarial decision aids over clinical decision-makers. And yet, the studies described above suggest that there may be something innately uncomfortable about trusting a computer that compels us instead to defer to expert human judges. Of course, trust in expert judges is by no means unfounded. Human judges are often very successful based solely on the powers of intuition, and experts can be trained to make complex judgments with great ease (Gladwell, 2005; Klein, 1998). However, the fact of the matter is simply that as medicine advances, difficult and ambiguous decisions about how best to treat a patient require physicians to make choices when it simply may be beyond their powers to make successfully.

Previous research has focused on overconfidence and patient discomfort with the usage of decision aids as sources of physician reluctance to utilize decision aids in practice. The possibility that physicians maintain a loss framework that makes the decision aid even less attractive corroborates many of these phenomena and provides an explanation of the behavior in cases where it was once lacking. For example, previous research simply stated that physicians feared derogation by patients (Arkes, Shaffer, and Medow, 2007; Cruickshank, 1985). However, research with mock jurors also showed that utilization of a decision aid was protective in cases in which a medical error had been made (Arkes, Shaffer and Medow, 2008). Prospect Theory-like

effects provide an explanation as to why physicians would risk more severe derogation in a malpractice suit by not using a decision aid in their practice. A risk-seeking attitude encourages physicians to forfeit the additional protection afforded by decision aids against the possibility that the physician will not make an errors.

Finally, it is true that actuarial decision aids have been controversial in the research literature. While many decision aids are helpful and improve clinical treatment, studies have shown that some do not (Garg et al. 2005; Hunt et al., 1998). We still lack substantial information about how to develop actuarial formulas that physicians will employ. Some decision aids do not improve medical care, whereas some do improve judgment but remain unappealing to physicians – in which case patient care suffers. Interventions such as calibration feedback for the physician (Arkes et al. 1986), and even to educate people to improve attitudes toward statistics have been suggested to increase usage of actuarial decision rules (Sieck and Arkes, 2005). Prospect theory suggests yet another possibility - that future work involving the framing of actuarial decision aids in a gain framework may be helpful in changing physician attitudes toward the actuarial decision aids in medicine.

References

Arkes, H. R., Dawes, R. M., & Christensen, C. (1986). Factors influencing the use of a decision rule in a probabilistic task. *Organizational Behavior and Human Decision Processes, 37*(1), 93-110. doi: 10.1016/0749-5978(86)90046-4

Arkes, H. R., Shaffer, V. A., & Medow, M. A. (2007). Patients derogate physicians who use a computer-assisted diagnostic aid. *Medical Decision Making, 27*(2), 189-202. doi:10.1177/0272989X06297391

Arkes, H. R., Shaffer, V. A., & Medow, M. A. (2008). The influence of a physician's use of a diagnostic decision aid on the malpractice verdicts of mock jurors. *Medical Decision Making, 28*(2), 201-208. doi:10.1177/0272989X07313280

Armor, D. A., & Taylor, S. E. (2002). When predictions fail: The dilemma of unrealistic optimism. *Heuristics and biases: The psychology of intuitive judgment.* (pp. 334-347) New York, NY, US: Cambridge University Press.

Bouaud, J., Seroussi, B., Antoine, E. C., Gozy, M., Khayat, D., & Boisvieux, J. F. (1998). Hypertextual navigation operationalizing generic clinical practice guidelines for patient-specific therapeutic decisions. *Proc AMIA Symp.,* , 488–492.

Cruickshank, P. J. (1985). Patient rating of doctors using computers. *Social Science & Medicine, 21*(6), 615-622. doi: 10.1016/0277-9536(85)90200-X

Dawes, R., Faust, D., & Meehl, P. (1989). Clinical versus actuarial judgment. *Science, 243*(4899), 1668-1674. doi:10.1126/science.2648573

De Dombal, F. T., Leaper, D. J., Horrocks, J. C., Staniland, J. R., & McCann, A. P. (1974). Human and computer-aided diagnosis of abdominal pain: Further report with emphasis on performance of clinicians. *British Medical Journal, 1*(5904), 376-380. doi:10.1136/bmj.1.5904.376

Evans, J. S. B. T. (2003). In two minds: Dual-process accounts of reasoning. *Trends in Cognitive Sciences, 7*(10), 454-459. doi: 10.1016/j.tics.2003.08.012

Garg, A. X., Adhikari, N. K. J., McDonald, H., Rosas-Arellano, M. P., Devereaux, P. J., Beyene, J., Sam, J., & Haynes, R. B. (2005). Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: A systematic review. *JAMA: The Journal of the American Medical Association, 293*(10), 1223-1238. doi:10.1001/jama.293.10.1223

Gladwell, M. (2005). *Blink: The power of thinking without thinking*. New York: Little, Brown ad Company.

Graham, I. D., Stiell, I. G., Laupacis, A., McAuley, L., Howell, M., Clancy, M., Durieux, P., Simon, N., Emparanza, J. I., Aginaga, J. R., O'Connor, A., & Wells, G. (2001). Awareness and use of the Ottawa ankle and knee rules in 5 countries: Can publication alone be enough to change practice? *Annals of Emergency Medicine, 37*(3), 259-266. doi: 10.1067/mem.2001.113506

Hunt, D. L., Haynes, R. B., Hanna, S. E., & Smith, K. (1998). Effects of computer-based clinical decision support systems on physician performance and patient outcomes: A systematic

review. *JAMA: The Journal of the American Medical Association, 280*(15), 1339-1346. doi:10.1001/jama.280.15.1339

Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist, 39*(4), 341-350. doi:10.1037/0003-066X.39.4.341

Kaplan, B. (2001). Evaluating informatics applications—clinical decision support systems literature review. *International Journal of Medical Informatics, 64*(1), 15-37. doi: 10.1016/S1386-5056(01)00183-6

Klein, G. (1998). *Sources of power: How people make decisions*. Cambridge, Massachusetts: The MIT Press.

Komaroff, A. L. (1982). *Algorithms and the 'art' of medicine* American Public Health Association.

McGlynn, E. A., Asch, S. M., Adams, J., Keesey, J., Hicks, J., DeCristofaro, A., & Kerr, E. A. (2003). The quality of health care delivered to adults in the united states. *The New England Journal of Medicine, 348*(26), 2635-2645. doi:10.1056/NEJMsa022615

Onkal, D., Goodwin, P., Thomson, M., Gonu, S., & Pollock, A. (2009). The relative influence of advice from human experts and statistical methods on forecast adjustments. *Journal of Behavioral Decision Making, 22*, 390-391-409.

Pezzo, M. V., & Pezzo, S. P. (2006). Physician evaluation after medical errors: Does having a computer decision aid help or hurt in hindsight? *Medical Decision Making, 26*(1), 48-56. doi:10.1177/0272989X05282644

Promberger, M., & Baron, J. (2006). Do patients trust computers? *Journal of Behavioral Decision Making, 19*(5), 455.

Redelmeier, D. A., & Shafir, E. (1995). Medical decision making in situations that offer multiple alternatives. *JAMA: The Journal of the American Medical Association, 273*(4), 302-305. doi:10.1001/jama.1995.03520280048038

Shen, O., Rabinowitz, R., Geist, R. R., & Shafir, E. (2010). Effect of background case characteristics on decisions in the delivery room. *Medical Decision Making,* doi:10.1177/0272989X09353451

Shrestha LB. (2006). *Life expectancy in the united states* No. CRS Report RL32792)US Congressional Research Service.

Sieck, W. R., & Arkes, H. R. (2005). The recalcitrance of overconfidence and its contribution to decision aid neglect. *Journal of Behavioral Decision Making, 18*(1), 29.

Spring, B. (2008). Health decision making: Lynchpin of evidence-based practice. *Medical Decision Making, 28*(6), 866-874. doi:10.1177/0272989X08326146

Swets, J. A., Dawes, R. M., & Monahan, J. (2000). Psychological science can improve diagnostic decisions. *Psychological Science in the Public Interest, 1*(1), 1-26.

Wade, T. C., & Baker, T. B. (1977). Opinions and use of psychological tests: A survey of clinical psychologists. *American Psychologist, 32*(10), 874-882. doi:10.1037/0003-066X.32.10.874

Wendt, T., Knaup-Gregori, P., & Winter, A. (2000). Decision support in medicine: A survey of problems of user acceptance. *Stud Health Technol. Inform., 77*, 852-6.