Health Care Management Science 7, 207–215, 2004

Stochastic Dominance and Medical Decision Making

MOSHE LESHNO *
Faculty of Management and Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, 69978, Israel
E-mail: leshnom@post.tau.ac.il

HAIM LEVY **
School of Business Administration, The Hebrew University, Jerusalem, 91905, Israel

Abstract. Stochastic Dominance (SD) criteria are decision making tools which allow us to choose among various strategies with only partial information on the decision makers' preferences. The notion of Stochastic Dominance has been extensively employed and developed in the area of economics, finance, agriculture, statistics, marketing and operation research since the late 1960s. For example, it may tell us which of two medical treatments with uncertain outcomes is preferred in the absence of full information on the patients' preferences. This paper presents a short review of the SD paradigm and demonstrates how the SD criteria may be employed in medical decision making, using the case of small abdominal aortic aneurysms as an illustration. Thus, for instance by assuming risk aversion one can employ second-degree stochastic dominance to divide the set of all possible treatments into the efficient set, from which the decision makers should always choose, and the inefficient (inferior) set. By employing Prospect Stochastic Dominance (PSD) a similar division can be conducted corresponding to all S-shaped utility functions.

Keywords: Stochastic Dominance, medical decision making, Prospect Stochastic Dominance

1. Introduction

We suggest in this paper that the Stochastic Dominance (SD) criteria that are generally employed in finance and economics may be even more relevant to medical decision making. Let us first demonstrate the application of SD criteria in financial decision making. Suppose that an investor has to choose one of n possible investments (e.g., stocks or portfolios). For various sets of partial information on preferences, SD rules are employed to divide all possible investment portfolios into two sets: the efficient set and the inefficient set. We can safely assert that all investments included in the inefficient set are inferior and the investor should select her investment portfolio from the efficient set.

Let us illustrate the SD concept with the most intuitive First-degree Stochastic Dominance (FSD), which corresponds to the trivial partial information assertion that people prefer more money to less money (i.e., the utility function is non-decreasing). Consider, for example only three investments, F, G and H, where F yields \$2000 or \$4000 with equal probability, G yields \$2500 or \$4000 with equal probability, and H yields \$1500 or \$7000 with equal probability. Applying the FSD rule, no investor who prefers more to less money will invest in F. However, some investors may prefer H over G and some may prefer G over H. Therefore, in this specific example, F is included in the inefficient set while G and H are included in the efficient set. Obviously, this is a trivial example and one does not need to be an expert to reach this conclusion. However, with many possible outcomes the selection is much more difficult. Moreover, as we shall see in this paper, there are other SD criteria which are more complicated and less intuitive. The advantage of the SD framework is that very little information on preferences is assumed and any assumptions that are made are general in nature. Moreover, there is no need to assume some specific statistical distribution of outcomes, such as the normal distribution. In this sense SD is a set of distribution-free decision rules. The disadvantage of SD rules in finance is that they cannot provide a recipe on how to diversify optimally among various assets (i.e., how to construct a portfolio out of the n available individual stocks). This disadvantage is irrelevant to medical decision making. To see this, suppose that five possible treatments (e.g., surgical, non-surgical, etc.) are considered. Each treatment implies a survival distribution (e.g., $0, 1, 2, \ldots, n$ years, with corresponding probabilities). Assuming that people always prefer to live more than less years (with the same life quality), we can employ SD decision rules to divide the above set of five possible treatment into, say, two which are included in the efficient set and three which are included in the inefficient set. The infinite set of distributions which exist in investment in stocks, that is, the infinite number of diversification policies for even a finite number of possible stocks, is irrelevant in medical decision making, insofar as mixing the strategies (e.g., surgery and no-surgery) is generally irrelevant. Therefore, applying the SD rule to medical decision making is even more relevant than applying it to a choice

* Corresponding author.
** Haim Levy acknowledges the financial support of Krueger Center of Finance.

among financial assets, as the diversification drawback of the SD rule is generally irrelevant in medical decision making.

The concept of stochastic dominance is not new. Karamata proved a theorem in 1932, for equal mean distributions which is very similar to second-degree stochastic dominance (SSD) [1–5]. The application of SD concepts in decision theory began about 45 years ago (see [6,7]), but the development of the theory of SD and its many theoretical and empirical extensions in economics and finance took place only after 1969–1970, when three papers were independently published by Hadar and Russell [8], Hanoch and Levy [9], Rothschild and Stiglitz [10] and a follow up paper was published by Whitmore [11]. Since then the concept of SD has been extensively employed in the area of economics, finance, agriculture, statistics, marketing, and operation research. For a survey and detailed elaboration of the various SD rules and their applications in various area, see [12–14]. We illustrate in this paper the possible application of the SD method to medical decision making with some artificial examples and with one set of data provided to us on a study recently published in the medical literature. We would like to stress at the outset that this paper focuses on medical decision making and not on clinical trials or the SD analysis of clinical trials. In applying SD to real medical data we assume that the distributions, which are based on clinical trials, are the population distributions and we do not attempt to conduct statistical significance tests. We only use the data to illustrate the use of the SD rules and how they can be employed in medical decision making. As in most medical decision making, the distributions over the outcomes are based on empirical studies such as clinical trials; in practice, statistical significance tests are needed. However, in this study we assume that the sample distribution is the population distribution and we make no statistical inference of SD results, an issue that has been exhaustively dealt with in a recent article by Barett and Donald [15].

While the SD rule has been applied to many research areas, to the best of our knowledge this is the first attempt to apply it to medical studies, where in our opinion its use is actually called for. However, as it is possible that many researchers in medicine are unaware of these decision making criteria, this paper presents a short review of the SD paradigm and demonstrates how one can employ SD criteria in medical decision making. Though we illustrate the use of SD criteria with data provided by a recently published study, the medical domain contains ample data to which SD can usefully and quite easily be applied.

The paper is organized as follows: section 2 presents a brief description of the stochastic dominance rules formulated in terms of the cumulative probability function. Section 3 is devoted to illustrating the application of the SD criteria to the case of small abdominal aortic aneurysms (AAA). Section 4 concludes the paper.

2. Decision making under uncertainty: stochastic dominance criteria

2.1. Expected utility theory

The expected utility theorem developed by von Neumann and Morgenstern [16] asserts that given three assumption (axioms) on the rational behaviour of decision makers, for each decision maker a numerical value is assigned to the various outcomes (called utility) such that for two decision alternatives the decision maker prefers the alternative with the higher *expected utility*, that is, the higher $Eu(x)$, where x is the outcome, u the utility and E the expected value operator.[1] In finance and economics, x is the financial outcome in dollars and $u(x)$ is the utility derived from x. In medicine the outcomes are generally measured in years of life adjusted for quality.[2] In this paper, we consider medical outcomes as length of life, assuming equality of quality of life. We denote by x staying alive exactly x years (months or weeks) and then dying. The Quality Adjusted Life Year (QALY) measure could equally well be used for the SD analysis. We can also assume that x may be discounted; thus, living for six QALYs would yield x^* QALYs < 6 due to discounting, where x^* is the discounted value. However, it can be proved that if one strategy is inferior to another by SD criteria without discounting, it will also be inferior under the discounted variable (for the sake of brevity, the proof, which is rather long, is omitted).

Decision makers may have different utility functions, depending on their preferences. Take for example two utility functions, $u_1(x) = x$ and $u_2(x) = \sqrt{x}$, where x stands for staying alive exactly x years and then dying. For the sake of simplicity, assume that the patient has two alternative treatment options: "early surgery" or "no early surgery", the quality of life being the same in either case. Moreover, suppose that early surgery implies death immediately with a probability of 4% and death after 5 years with a probability of 96%, whereas no early surgery implies death with certainty after 4.7 years. What should the patient choose? Though early surgery increases life expectancy from 4.7 years to 4.8 years ($4\% \times 0$ years $+ 96\% \times 5$ years $= 4.8$ years), there is no question that some proportion of the patients will choose not to have early surgery due to the slight risk of dying immediately. In terms of utility, the expected utility derived from living "0" with a 4% chance or 5 years with 96% may be lower than the utility associated ith living exactly 4.7 years for sure. Thus, in principle, to make an optimal decision, which may vary from one patient to another, one needs to know the patient's utility with its inherent preferences. Consider our two decision makers, each with one of the two utility functions defined above, facing the two given treatment alternatives. The expected utility for the first de-

[1] A formal presentation of the expected utility theorem can be found in [7,16,17]. We should note parenthetically that the original paper by von Neumann and Morgenstern assumes nine assumptions; however three assumptions have been shown to be enough.

[2] A generalization of SD to the case where the outcomes are only ordinal was developed by Spector et al. [18].

cision maker ($u(x) = x$) having "early surgery" is given by $4\% \times 0 + 96\% \times 5 = 4.8$, compared to the expected utility for "no early surgery", which is given by $1 \times 4.7 = 4.7$. This specific decision maker prefers "early surgery" over "no early surgery". However, for the second decision maker with the utility function u_2, we have that the expected utility induced by "early surgery" is $4\% \times \sqrt{0} + 96\% \times \sqrt{5} = 2.14$, compared to "no early surgery", where the expected utility is given by $1 \times \sqrt{4.7} = 2.16$. Thus the second decision maker prefers "no early surgery" over "early surgery". The difference between these two hypothetical decision makers is that the first is neutral towards risk whereas the second is risk averse.

In the above calculation we assume that no cost is involved. However, as the decision generally also involves a cost, the utility function is two-dimensional: $u(x, c)$, where x is number of years lived at c cost. Discounting x and c should also be taken into account in the expected utility maximization. In this paper we focus on stochastic dominance rules applying to x as the single random variable. However, the analysis can be extended to two-dimensional stochastic dominance where both x and c are random variables (see [19]).

2.2. *Risk and optimal treatment*

Most economic models assume risk aversion, that is, diminishing marginal utility of money. Though not all agree with this assumption, something that may be conceptually very different from the utility of money may be relevant to this paper: the utility of living an additional year. Therefore, in this section we discuss the utility of living x years. Pliskin et al. used multiattribute utility theory to suggest a form for a utility function over life years and health status [20]. What do we know about the general characteristic of such a utility function? McNeil et al. [21] found that most patients are risk averse, that is, the utility from each additional year of living is positive, but it increases at a diminishing rate.

Verhoef et al. [22], found empirically that most patients have an S-shaped preference curve as suggested by Kahneman and Tversky's Prospect Theory (PT) [23], where the inflection point of the utility is at some aspiration level of years to live, for example, the number of years between the patient's current age and her age-adjusted life expectancy. Nease [24] showed that the utility function can take various shapes, the S-shaped being only one possible form. Moreover, he concluded that:

"*In summary, the work of Verhoef and colleagues provokes us to consider afresh the assumption of risk neutrality we often make in our analyses, to use caution in applying practice guidelines that fail to account for the preferences of the individual patient*" (see [24, p. 203]).

Though there is disagreement on the shape of the utility function almost all agree that the patients are rational, that is, they prefer to live a longer rather than a shorter life,[3] in which case one of the SD criteria (see below) applies.

Obviously, medical researchers are well aware of the fact that a patient's utility from each year of living should be a factor in choosing the optimal strategy. Yet, in choosing the medical decision that maximizes life expectancy, one implicitly assumes that $u(x) = x$, that is, a linear utility function or risk neutrality. If preferences are so important why then is it so common to rely on life expectancy, that is, to assume risk neutrality in the decision making process? Presumably, there are two reasons. (a) The main reason for the risk-neutrality (or a linear utility) assumption is the fact that the patient's utility function is unknown. (b) As we shall see below, the strategy with the highest expected value is always an efficient strategy in the SD framework; hence the commonly employed criterion can never be inferior for all patients. Nevertheless, in this paper we suggest an optimal decision making framework regarding two (or more) possible medical treatments (e.g., early surgery, radiation or surveillance) even when we have only general information on preferences. In other words, establishing the efficient set of strategies is based on partial information regarding the patient's preferences. The SD methodology is based on expected utility, which incorporates risk into the analysis. Though the strategy with the highest life expectancy is always included in the efficient set, other strategies may also be included, and most patients may even prefer them over the highest life expectancy strategy. Sometimes, there is a dominance between two treatments in which case the SD rules can be very usefully employed by physicians in choosing between two treatment strategies. In cases where there is no dominance between two alternative treatments, what remains is to present the information to the patient and let her make the choice, or, if this is impossible, it should be left to the physician and the family to make the choice.

2.3. *Stochastic dominance rules*

In this section we provide the stochastic dominance rules, each of which depends on a different level of partial information on preferences. We define the various decision rules, providing a numerical example and an intuitive explanation.

2.3.1. *First-degree stochastic dominance (FSD)*

We define as x the real number of years a patient lives until she dies; for example, $x = 10.5$ means she lives exactly 10.5 years and then dies. The cumulative probability is defined by $F(x) = \Pr(X \leqslant x)$, where Pr stands for probability and $F(x)$ is the probability of staying alive no longer than x years. The goal is to maximize the patient's expected utility, considering all possible strategies.

Suppose that there are two mortality distributions[4] F and G, which are induced by two treatment strategies. We

[3] Some studies have measured utilities for health states that are worse than death (e.g., [25,26]). In this study we use QALY i.e., number of years lived adjusted for quality. Therefore, a very bad year from the patient's point of view can be counted as "zero" if the suffering is too high.

[4] The same analysis can be conducted with survival curves instead of mortality distributions, where given a mortality distribution F the corresponding survival curve is given by $S = 1 - F$.

say that F dominates G by first-degree stochastic dominance (FSD) if and only if,

$$F(x) \leqslant G(x), \quad \text{for all } x, \text{ and} \qquad F(x_0) < G(x_0), \quad \text{for some } x_0. \tag{1}$$

A utility which is non-decreasing in years means that the patients always prefer to live up to year $x + \varepsilon$ rather than up to year x for all x and $\varepsilon > 0$. This means that in such cases we can make a choice even without knowing the precise utility function. The only assumption is that the more years the patient lives, the better off she is. If F dominates G by FSD we can safely assert that $E_F(u(x)) \geqslant E_G(u(x))$, for all non-decreasing utility functions u, where $E_F(u(x))$ and $E_G(u(x))$ stand for the expected utility of u under the F and G distributions, respectively. Thus, if equation (1) holds, treatment F dominates treatment G for all types of utility function, as long as they are non-decreasing in years.

2.3.2. *Numerical example and intuitive explanation of FSD*

Suppose that two alternative medical treatments denoted by F and G are associated with the following mortality distributions:

G		*F*	
x (years)	Probability of *x*	*x* (years)	Probability of *x*
1	0.01	1	0
2	0.20	2	0.10
3	0.40	3	0.50
4	0.29	4	0.30
5	0.10	5	0.05
–	–	6	0.05

What should the patient choose? It is not at all obvious from the above table. However, drawing the cumulative distribution of F and G (see figure 1) reveals that $F(x) \leqslant G(x)$ for all x; hence F is below G and therefore F is better than G for all preferences.

The intuitive explanation for why $F(x) \leqslant G(x)$ implies that F is preferred under all patient utility preferences is not straightforward. However, recall that $F(x) \leqslant G(x)$ implies that $1 - F(x) \geqslant 1 - G(x)$. And as $1 - F(x)$ is the probability of living x years *or more* under F, the fact that F is below G implies that under decision F the probability of living x years or more is greater than the probability corresponding to decision G, certainly a property desired by all patients with a non-decreasing utility function. And as this probability is greater for every selected value of x, this explains why F dominates G for all preferences.

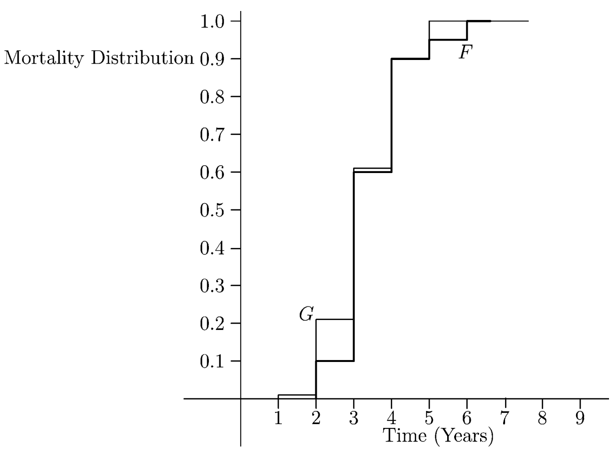

Figure 1. Mortality distributions: F dominates G by FSD.

The problem with FSD is that it is very unlikely that the mortality distribution corresponding to two decisions do not cross. And if they cross there is no FSD dominance. Hence, the next step is to add an assumption regarding preferences such that a sharper decision is obtained even when the distributions under consideration cross.

2.3.3. *Second-degree stochastic dominance (SSD)*

Suppose that F and G cross, and hence there is no FSD. If we are now willing to assume risk-averse utility functions (i.e., risk aversion implying that each additional year one lives adds to the utility but at a diminishing rate) we may find a dominance even if F and G cross. The decision criterion corresponding to diminishing marginal utility is called Second-degree Stochastic Dominance (SSD), as we make an assumption regarding the second derivative of u. That is, not only is $u' \geqslant 0$ but also $u'' \leqslant 0$.

F dominates G by Second-degree Stochastic Dominance (SSD) if and only if

$$\int_0^x \big[G(t) - F(t)\big]\mathrm{d}t \geqslant 0, \quad \text{for all } x, \text{ and there is a strict inequality for some } x_0. \tag{2}$$

If equation (2) holds it implies that for risk averters, treatment F is better. Formally, it implies that $E_F(u(x)) \geqslant E_G(u(x))$ for all non-decreasing concave utility functions.

2.3.4. *Numerical example and intuitive explanation of SSD*

Let us assume the following two simple mortality distributions corresponding to two alternative medical decisions, F and G:

F		*G*	
x (years)	Probability of *x*	*x* (years)	Probability of *x*
2	0.50	1	0.50
4	0.50	5	0.50

First note that the two cumulative distributions cross (see figure 2); hence neither F nor G dominates the other by FSD. However, as figure 2 reveals $\int_0^x [G(t) - F(t)]\mathrm{d}t \geqslant 0$ for all x,[5] and therefore we can safely conclude that for all risk averters (i.e., u with $u' > 0$ and $u'' < 0$), F dominates G.

[5] Note that for $1 \leqslant t \leqslant 2$, $[G(t) - F(t)] = \frac{1}{2}$ and for $4 \leqslant t \leqslant 5$, $[G(t) - F(t)] = -\frac{1}{2}$; therefore $\int_1^2 [G(t) - F(t)] = \frac{1}{2}$ and $\int_4^5 [G(t) - F(t)] = -\frac{1}{2}$, and thus $\int_0^x [G(t) - F(t)]\mathrm{d}t \geqslant 0$ for all x.

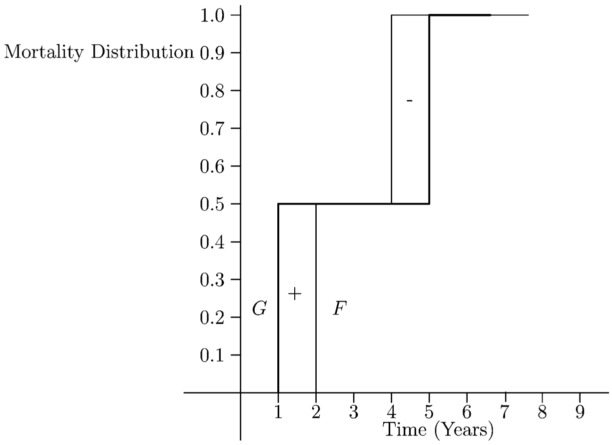

Figure 2. Mortality distributions: F dominates G by SSD but not by FSD.

To see the intuition of SSD let us look at the difference between F and G. If one shifts from G to F she earns one year of living (comparing 1 and 2) with a probability of 0.5 and she loses one year of living (comparing 5 and 4) with a probability of 0.5. As the marginal utility of living is decreasing, the gain from the shift from 1 to 2 is larger than the loss from shifting from 5 to 4; hence F dominates G for all risk averters. Take for example the utility function $u(x) = \sqrt{x}$. It is easy to see that $u' > 0$ and $u'' < 0$. Indeed for the utility function we have a gain from shifting from 1 to 2 of $\sqrt{2} - \sqrt{1} \cong 1.41 - 1 = 0.41$. The loss from shifting from 5 to 4 in years is $-\sqrt{5} + \sqrt{4} \cong -2.24 + 2 = -0.24$. As the marginal utility gain is larger than the marginal utility loss, for this risk averter a shift from G to F is worthwhile. As the same argument is true for all types of preferences with $u' > 0$ and $u'' < 0$, F dominates G by SSD.

Before we turn to Prospect Stochastic Dominance (PSD), note that with monetary outcomes Levy and Sarnat [27] show that the discounting factor can be ignored as dominance without the discounting factor in a two-period scenario implies a dominance with any positive discount factor. The proof with x as the number of years to live is much more complicated, but it is valid and hence the discount factor can be ignored also in the case dealt with here.

2.3.5. Prospect stochastic dominance (PSD)

While FSD and SSD were defined on all values of x without a separation between negative and positive outcomes, the PSD include two ranges of outcomes; the negative and the positive ranges. The reason for this separate treatment of the two ranges is the inflection point in the preferences. For money outcomes, $x = 0$ is the inflection point. In medicine, as the outcomes are number of years adjusted for quality it is reasonable to define the inflection point as the number of years one is expected to live, adjusted for quality, which is called an aspiration level. As already noted, Verhoef et al. [22] experimentally found that the utility function is S-shaped with an inflection point at the aspiration level. For example, if the life expectancy is 80 years, then it is reasonable to take the aspiration level as 80 years. However, even for a given aspiration level, no one knows the parameters of the S-shaped function. Moreover, it may vary from one patient to another. However, it has been proved [12,28], that for any S-shaped function with aspiration level Z_{AL}, treatment F dominates treatment G if and only if

$$\int_y^x \big[G(t) - F(t)\big]\mathrm{d}t \geqslant 0, \quad \text{for all } y < Z_{\mathrm{AL}} \text{ and for all } x > Z_{\mathrm{AL}}, \tag{3}$$

where Z_{AL} is the aspiration level. In monetary bets $Z_{\mathrm{AL}} = 0$ (see [23]). However, in medical research it could be the number of years that patients with a given illness are expected to live. Of course, it is meaningful to employ the PSD rule only to a group of patients with the same aspiration level.

2.3.6. Numerical example and intuitive explanation of PSD

Let F and G given below be two mortality distributions for two alternative medical decisions:

G		*F*	
x (years)	Probability of x	x (years)	Probability of x
3	0.25	2	0.25
4	0.25	5	0.25
6	0.25	7	0.25
9	0.25	8	0.25

Suppose that the aspiration level is $z_{\mathrm{AL}} = 5.5$ years for the group of patients for which these two treatments are being consider. Figure 3 shows the cumulative distributions of F and G. As can be seen, for all $y < 5.5$ we have $\int_y^{5.5}[G(t) - F(t)]\mathrm{d}t \geqslant 0$ and for all $x > 5.5$ we have $\int_{5.5}^x[G(t) - F(t)]\mathrm{d}t \geqslant 0$; hence F dominates G for all S-shaped preferences. The intuitive explanation for the range $x > z_{\mathrm{AL}} = 5.5$ is the same as for SSD, as risk aversion prevails in this range. For $x < z_{\mathrm{AL}} = 5.5$ the S-shaped function is characterized by a risk-seeking segment ($u' > 0$ and $u'' > 0$); hence moving from the pair (3, 4) to (2, 5) is desirable as the marginal utility is increasing and the loss due to

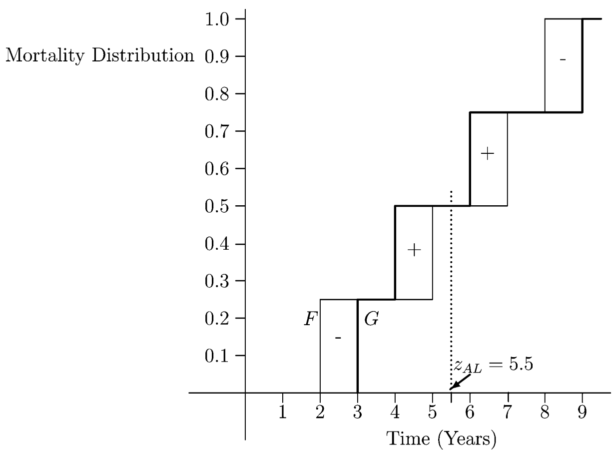

Figure 3. Mortality distributions: no FSD, no SSD but F dominates G by PSD.

the shift from 3 to 2 is thus smaller than the gain from shifting from 4 to 5. And for this reason, whenever equation (3) holds, $E_F u(X) \geqslant E_G u(X)$ for all S-shaped utility functions.

Figure 4 illustrates FSD, SSD and PSD. In figure 4(a), F dominates G by FSD as the two distributions do not cross. In figure 4(b) there is no FSD, but F dominates G by SSD. In figure 4(c) there is no FSD and no SSD; however, F dominates G by PSD, with an aspiration level denoted by Z_{AL}. Note that the PSD is sensitive to the aspiration level.

Thus, one can employ FSD, SSD and PSD to find the efficient set of mortality distributions (induced by various medical decisions). However, a necessary condition for dominance by all these rules is that the life expectancy of the superior mortality distribution is larger than or equal to the life expectancy of the inferior life mortality distribution (a proof can be found in [12]). Therefore, the medical decision corresponding to the highest life expectancy is always included in the efficient set, that is, no other mortality distribution can dominate it. We may however have other mortality distributions with lower means which are also efficient as no other distribution, not even the one with the highest life expectancy, dominates them. The fact that the distribution with the highest mean is always included in the efficient set may partially explain why the goal of maximizing life expectancy (or QALYs) is so popular.

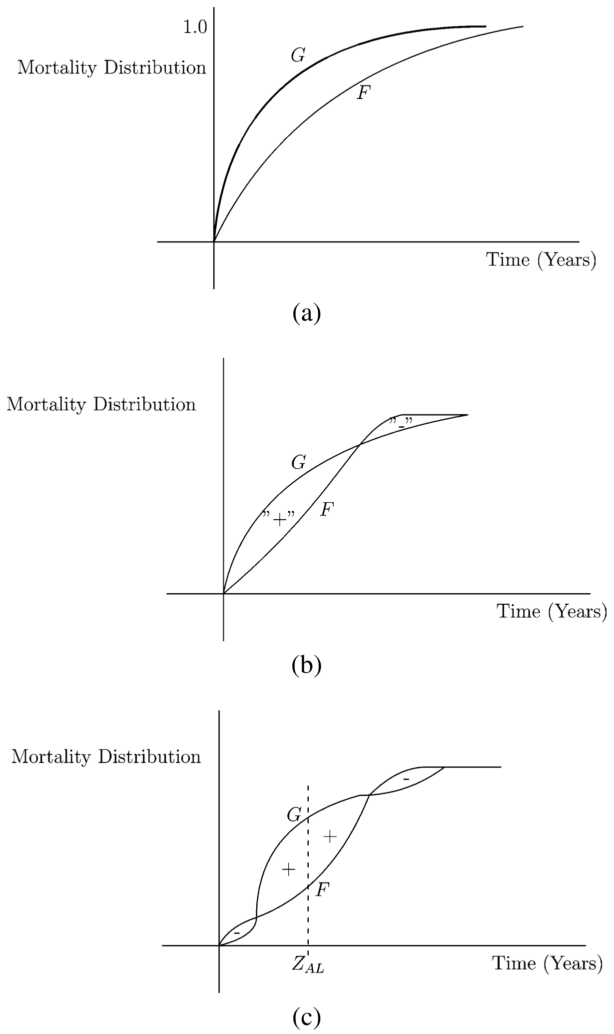

Figure 4. Two hypothetical mortality distributions, where (a) F dominates G by FSD, (b) F dominates G by SSD but not by FSD, (c) F dominates G by PSD but not by FSD or SSD.

2.3.7. "Almost" stochastic dominance

All stochastic dominance rules can be refined by employing stronger rules, the so-called "almost" stochastic dominance rules developed recently by Leshno and Levy [29]. Though we will not elaborate on these rules, let us just illustrate the concept by a simple example. Suppose that treatment A implies that $x_1 = 3$ years with probability 0.01 or $x_2 = 70$ years with probability 0.99. Treatment B implies 3.1 years with probability 1. It is easy to see that neither treatment dominates the other by FSD or SSD, but any given experiment may show "all" or "almost all" patients choosing treatment A. Leshno and Levy [29] developed a set of modified SD rules called "almost" SD, corresponding to "almost all" patients. "Almost" SD rules allow us to make a choice which cannot be made by the common SD rules as the two treatments are efficient. Thus, Leshno and Levy show that in the above example treatment B dominates treatment A despite the fact that the two cumulative distributions cross. However, in the rest of the paper we focus on the common SD rules, bearing in mind that these rules can be improved by "almost" SD rules.

The proofs of FSD, SSD and PSD as well as the necessary condition that the strategy with the highest mean is always efficient can be found in [12,28]. The proof of "almost" SD is given [29]. Let us now turn to illustrate SD criteria with real medical data.

3. Employing SD criteria in the small abdominal aortic aneurysms case

In this section we illustrate the application of the SD paradigm to actual medical data, derived from two studies recently published in the *New England Journal of Medicine* on the effect of surgery on patients with abdominal aortic aneurysms (AAA) 4.0 to 5.4 cm in diameter [30,31]. The mortality distributions of two groups were compared: the first group had immediate open surgical repair and the other group underwent surveillance only. In one study (the U.S. study, [30]) the authors conclude that the survival rate does not improve with elective surgical repair. The other study (the U.K. study, [31]) concludes that there is no significant long-term difference in the mean survival between the early surgery and surveillance groups, although after eight years total mortality was lower in the early surgery group. In the two studies the mortality probability distributions of the two groups are given graphically; however, the focus is on the *life expectancy.* In both studies the means of the years the patients lived are not different from each other.

In the U.K. study [31] the restricted mean duration of survival was 6.5 years among patients in the surveillance group, as compared to 6.7 years among patients in the early surgery

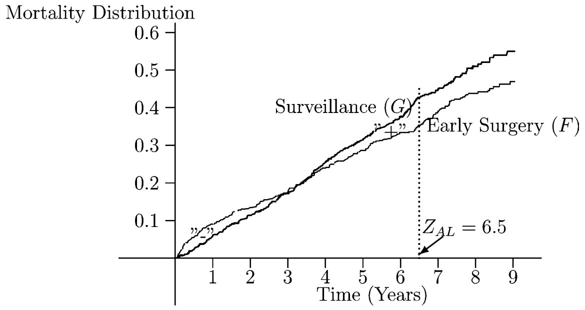

Figure 5. The cumulative mortality distributions of "early surgery" and "surveillance".

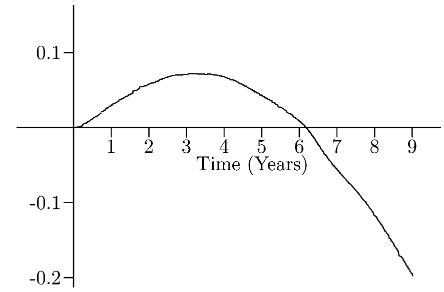

Figure 6. $T(x) = \int_0^x [F(t) - G(g)]\mathrm{d}t$ where F represents "early surgery" and G "surveillance".

group. Once again, these means are not statistically different. However, further evaluation after eight years revealed a small but statistically significant difference in mortality that favors early surgery. Taking the mortality distribution reported in this study as the population mortality distribution of patients with AAA, can we conclude based of the life expectancy figures from these two studies that early surgery should be avoided? Or should surgery be conducted? As we shall see, the most important feature of SD decision making rules is that they are based on the distributions of survival with and without the surgery and do not rely only on life expectancy. The interpretation of the result of this study is far from simple, despite the fact that the means of life duration in years of the two groups are not statistically different.[6] Short-term survival was worse in the early surgery group and longer-term survival was worse in the surveillance group. That is, the two survival distributions cross. Furthermore, suppose that these data are related to the population rather than to a sample, and hence no statistical issues arise. Suppose that by having early surgery, you increase the risk of death immediately from "0" to 2%, and you simultaneously decrease the chance of death in any of the next n years, such that the life expectancy is unchanged. Would you prefer early surgery? To complicate this question, what is your choice when the life expectancy due to early surgery increases by x years? The choice is fully dependent on the patient's preferences and in particular on the patient's attitude toward risk. That is, some patients may prefer to avoid surgery even if life expectancy increases as there is a small probability of death due to the surgery itself. We illustrate how to employ the suggested SD criteria on the U.K. study datal, which were kindly provided to us by Professor Janet Powell.

Figure 5 provides the cumulative distribution mortality functions F and G (where F is cumulative mortality corresponding to the early surgery group and G is the cumulative mortality probability function corresponding to the surveillance group. Because of right censoring the graph is truncated when it approaches 0.6). As we can see for small numbers of years, the probability of death is higher under F than G. However, for about $x \geqslant 3.6$, G is higher than F. Because F and G cross, there is no FSD dominance. Next we check whether there is SSD. To check whether F (or early surgery) dominates G (surveillance) we check whether the integral $\int_0^x [G(t) - F(t)]\mathrm{d}t$ is positive for all values of x. It is clearly negative for all $x < 3.6$ years (see figure 5). We also need to check whether G dominates F by SSD. We find that this is not the case, as the relevant integral $\int_0^x [F(t) - G(t)]\mathrm{d}t$ is negative for some value of x. Indeed figure 6, which shows the function $T(x) \equiv \int_0^x [F(t) - G(t)]\mathrm{d}t$, reveals that the curve $T(x)$ is negative for $x \geqslant 6.19$, implying that G does not dominate F by SSD either. Thus, neither F nor G dominates the other by FSD or SSD, and based on these rules we cannot tell whether early surgery is preferred or not.

Where do these results lead us? If one is willing to assume only that the utility of each additional year of living is positive, or that it is positive but with declining marginal utility, we conclude that early surgery is not necessarily an optimal policy for all patients. Surveillance is also not necessarily optimal. Actually, in this specific case, the efficient set includes both strategies and it is up to the patient to decide whether to undergo the operation because there is no dominating strategy. Some may express a preference for early surgery and some will not. Thus, the fact that early surgery increases life expectancy does not imply that it is optimal.

Now let us turn to the family of preferences suggested by Verhoef et al. We first need to define the aspiration life level, not a simple task, and check whether equation (3) holds. To employ PSD we need to find a group of patients with the same aspiration level, which can only be obtained from a careful elicitation of the individual patients' utility function. However, if, for the sake of argument only, we assume that the aspiration level is 6.5 years (because the mean duration of survival was 6.5 years), then, early surgery may be optimal. Looking at figure 5, we see that $\int_{6.5}^x [G(t) - F(t)]\mathrm{d}t \geqslant 0$ for all x because G is above F in this range. In the range $y < 6.5$ years we find that F and G cross each other several times. However, we find that $\int_y^{6.5} [G(t) - F(t)]\mathrm{d}t > 0$, for all $y < 6.5$ (we used Microsoft Excel to calculate it), hence F dominates G for all S-shaped functions and it is optimal

[6] Indeed, in the editorials which accompanied these two studies, R. Thompson (the editor of the journal) asserts: "Do the results of the Aneurysm Detection and Management (ADAM) study mean that no patient should undergo repair of an abdominal aortic aneurysm which is less than 5.5 cm in diameter? Unfortunately, the answer is not that simple" (see [32, p. 1484]). Thompson mentions several important factors (e.g., gender, smoking, etc.), and concludes that "it should be recognized that some carefully selected patients will still benefit from early surgery treatment" [32, p. 1485]. While Thompson relates to the health and other characteristics of the patient, in this study we discuss another facet of the patient's characteristics: her preferences (or utility), the risk attitude and the choice of which is the most beneficial to the patient.

to have early surgery for all patients with an S-shaped preference as long as they have the same 6.5 years aspiration level. Obviously, F also dominates G for a high value of Z_{AL} (see figure 5 and equation (3)). Thus, for $Z_{\mathrm{AL}} \geqslant 6.5$ years or more, early surgery is optimal. The minimum Z_{AL} (calculated from the data corresponding to figure 5) that guarantees PSD of F over G is 6.019. Thus, with these specific data there is no FSD and no SSD, but there is PSD for all $Z_{\mathrm{AL}} \geqslant 6.19$ years.

In summary, we found that neither early surgery nor surveillance dominates the other by FSD or SSD. Thus there are people who are risk averse who prefer early surgery and other people also risk averse who would prefer to undergo surveillance, and we cannot derive any conclusion that early surgery (or surveillance) is preferred by all people who are risk averse. Thus, based on FSD and SSD we cannot tell whether early surgery is preferred or not. However, based on PSD, we can conclude that all people with an S-shaped preference utility function and an aspiration level of 6.5 years (or more) would prefer early surgery over surveillance.

4. Summary and conclusion

Frequently in medical decision making one faces several possible treatments, each of which implies a given mortality distribution. The standard approach assumes (implicitly) that all subjects have the same preferences and utility function and thus the most common way is to choose the treatment with the highest life expectancy (with or without adjustment for quality). The assumption that all subjects have the same preferences and utility function is a fallacy. Thus, given two treatment options some subjects may prefer the first treatment and others may prefer the second. However, if we have more information on the preferences of a group of subjects (e.g., all subjects are risk avertors) Stochastic Dominance rules may reveal that the first treatment is preferred by all risk avertors. Thus, unlike the standard approach, SD rules provide a powerful tool to implement decisions that match the preferences of a group of people who may have different utility functions (e.g., risk averters).

In this paper we illustrate the SD concept in medical decision making. We illustrate the SD criteria with several simple numerical examples and with specific mortality tables.

For a given set of possible treatments and the corresponding mortality distributions, the SD rules divide the set of all possible treatments into efficient and inefficient treatments. The strategy which implies the highest life expectancy is always included in the efficient set of all SD rules. However, a strategy with a lower life expectancy may also be efficient (and hence may be desired by some patients) if it corresponds to, say, a lower probability of death immediately (as for example in the case where surgery is avoided). When we have FSD of two mortality distributions it is easy to recommend and suggest the superior treatment strategy. However, in many cases there is no FSD between two strategies and the SSD or PSD offer a chance of arriving at a definite conclusion and recommendation. The important feature in the SD paradigm is that whenever we get SD between two strategies we can determine the choice for a large group of decision makers (patients in our case) without the need to know each individual preference. If strategy F dominates G by FSD we can assert that all “rational” decision makers will prefer F over G. If strategy F dominates G by SSD we can assert that all decision makers who are risk averse will prefer F over G. Similarly, a PSD dominance holds for all S-shaped preferences.

Finally we would like to note the possible avenues of further research on SD and medical decision making. For example, one can employ two-dimensional SD by considering the cumulative distribution $F(x, c)$, where x is years lived (or QALYs) and c is the cost involved (both random variables). Also, discounting of these two variables can be incorporated into cost-effectiveness analysis. Finally, though in many cases we fail to get dominance of FSD, SSD or even PSD, due to a small violation of their rules, it is quite intuitive that “most” decision makers will prefer strategy F over strategy G. Recently, Leshno and Levy [29] developed the Approximately Stochastic Dominance criteria (ASD), according to which when F does not dominate G by FSD, some minor correction of F or G may yield dominance and therefore we can conclude that “most” decision makers will prefer F over G. The application of ASD rules to medicine and the SD paradigm to cost effectiveness analysis and other areas will be dealt with in a separate paper.

References

[1] J. Karamata, Sur une inegalite relative aux fonctions convexes, Publ. Math. Univ. de Belgrade 1 (1932) 145–148.

[2] M. Lebreton, Stochastic dominance: A bibliographical rectification and restatement of Whitmore’s theorem, Math. Social. Sci. 13 (1987) 73–79.

[3] J.E. Littlewood, G.H. Hardy and G. Polya, *Inequalities* (University Press, Cambridge, MA, 1959).

[4] S. Sherman, On a theorem of Hardy, Littlewood, Polya, and Blackwell, Proc. Nat. Acad. Sci. 37 (1951) 826–831.

[5] E. Lehmann, Ordered families of distributions, Ann. Math. Statist. 26 (1955) 399–419.

[6] P. Quirk and R. Saposnik, Admissibility and measurement utility function, Review of Economic Studies 29 (1962) 140–146.

[7] P.C. Fishburn, *Decision and Values Theory* (Wiley, NY, 1964).

[8] J. Hadar and W.R. Russell, Rules for ordering uncertain prospects, American Economic Review 59 (1969) 25–34.

[9] G. Hanoch and H. Levy, The efficiency analysis of choices involving risk, Review of Economic Studies 36 (1969) 335–346.

[10] M. Rothschild and J.E. Stiglitz, Increasing risk: I. Definition, Journal of Economic Theory 2 (1970) 225–243.

[11] G.A. Whitmore, Third degree stochastic dominance, Amer. Economic. Rev. 60 (1970) 457–459.

[12] H. Levy, *Stochastic Dominance: Investment Decision Making Under Uncertainty* (Kluwer Academic, Boston, 1998).

[13] H. Levy, Stochastic dominance and expected utility: Survey and analysis, Management Science (1992) 555–593.

[14] H. Levy and M. Leshno, Decision-making under uncertainty: Stochastic dominance, in: *Current State of Business Disciplines*, ed. S.B. Dahiya (Spellbound Publication, India, 2000) pp. 1377–1408.

[15] G.F. Barett and G. Donald, Consistent tests for stochastic dominance, Econometrica 71 (2003) 71–104.

[16] J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior* (Princeton University Press, Princeton, NJ, 1947).
[17] B.J. Cohen, Is expected utility theory normative for medical decision making? Medical Decision Making 16 (1996) 1–6.
[18] Y. Spector, M. Leshno and M. Ben-Horin, Stochastic dominance in an ordinal world, European Journal of Operational Research 93 (1996) 620–626.
[19] H. Levy and J. Paroush, Toward multivariate efficiency criteria, Journal of Economic Theory 7 (1975) 129–142.
[20] J. Pliskin, D.S. Shepard and M.C. Weinstein, Utility functions for life years and health status, Operations Researh 28 (1980) 206–224.
[21] B.J. McNeil, R. Weichselbaum and S.G. Pauker, Fallacy of the five-year survival in lung cancer, New England Journal of Medicine 299 (1978) 1397–1401.
[22] L.C.G. Verhoef, A.F.J. De Haan and W.A.J. Van Daal, Risk attitudes in gambles involving length of life: Empirical support for prospect theory, Medical Decision Making 14 (1994) 194–200.
[23] D. Kahneman and A. Tversky, Prospect theory: An analysis of decision under risk, Econometrica 47 (1979) 263–291.
[24] R.F. Nease, Risk attitudes in gambles involving length of life: Aspiration, variation and ruminations, Medical Decision Making 14 (1994) 201–203.
[25] D.L. Patrick, H.E. Starks, K.C. Cain, R.F. Uhlmann and R.A. Pearlman, Measuring preferences for health states worse than death, Medical Decision Making 14 (1994) 9–18.
[26] D. Feeny, W. Furlong, G.W. Torrence, C.H. Goldsmith, Z. Zhu, S. DePauw, M. Denton and M. Boyle, Multiattribute and single-attribute utility functions for the health utilities index mark 3 system, Medical Care 40 (2002) 113–128.
[27] H. Levy and M. Sarnat, Two-period selection and investors' discount rates, Journal of Finance 26 (1971) 757–761.
[28] H. Levy and Z. Wiener, Stochastic dominance and prospect dominance with subjective weighting functions, Journal of Risk and Uncertainty 16(2) (1998) 147–163.
[29] M. Leshno and H. Levy, Preferred by all and preferred by most decision makers: Almost stochastic dominance, Management Science 48 (2002) 1074–1085.
[30] F. Lederle, S. Wilson, G. Johnson, D. Reinke, F. Littooy, C. Acher, D. Ballard, L. Messina, I. Gordon, E. Chute, W. Krupski and D. Bandyk for The Aneurysm Detection, and Management Veterans Affairs Cooperative Study Group, Immediate repair compared with surveillance of small abdominal aortic aneurysms, New England Journal of Medicine 346 (2002) 1437–1444.
[31] The United Kingdom Small Aneurysm Trial Participants, Long-term outcomes of immediate repair compared with surveillance of small abdominal aortic aneurysms, New England Journal of Medicine 346 (2002) 1445–1452.
[32] R.W. Thompson, Detection and management of small aortic aneurysms, New England Journal of Medicine 346 (2002) 1484–1485.