A Decision-Theoretic Method for Analyzing Crossing Survival Curves in Healthcare

Appelbaum, Elie and Leshno, Moshe and Prisman, Eitan and Prisman, Eliezer, Z.

York University, Tel Aviv University, University of British Columbia, York University

20 March 2025

A Decision-Theoretic Method for Analyzing Crossing Survival Curves in Healthcare*

Elie Appelbaum†, Moshe Leshno‡, Eitan Prisman§, Eliezer Z. Prisman¶

May 9, 2025

Abstract

The problem of crossing Kaplan-Meier curves has not been solved in the medical research literature to date. This paper integrates survival curve comparisons into decision theory, providing a theoretical framework and a solution to the problem of crossing Kaplan-Meier curves. The application of decision theory allows us to apply stochastic dominance concepts and risk preference attributes to compare treatments even when standard Kaplan-Meier curves cross. The paper shows that as additional risk preference attributes are adopted, Kaplan-Meier curves can be ranked under weaker restrictions, namely with higher orders of stochastic dominance. Consequently, even Kaplan-Meier curves that cross may be ranked. The method we present allows us to extract all possible information from survival functions; hence, superior treatments that cannot be identified using standard Kaplan-Meier curves may become identifiable. Our methodology is applied to two examples of published empirical medical studies. We show that treatments deemed non-comparable because their Kaplan-Meier curves intersect can be compared using our method.

KEYWORDS: Survival Curve Analysis; Decision Theory; Risk Preference Modeling; Stochastic Dominance; Medical Treatment Comparison; Healthcare Data Interpretation

*We wish to thank Andrea Messori for his invaluable help and the data he provided. We also wish to thank the editor, associate editors and referees for their helpful comments and suggestions.

†Department of Economics, York University, Toronto, Canada

‡Faculty of medicine, Tel Aviv university, Sourasky Tel Aviv medical center and Coller school of Management, Tel Aviv university

§Vancouver General Hospital, Diamond Health Care Centre, & University of British Columbia, Vancouver, Canada

¶Schulich School of Business, York University, Toronto, Canada

1 Introduction

In their seminal paper, Kaplan and Meier (1958) provided what became known as the Kaplan-Meier estimator or the Kaplan-Meier Curve (KMC). These curves are non-parametric estimators of underlying theoretical survival functions (SFs), which, in turn, are simply statistical complementary cumulative distribution functions. Hence, the KMCs are estimators of such functions. Over the years, Kaplan-Meier Curves (KMCs) have become the gold standard, used extensively in countless applications in various medical (and other) fields to compare survival (or, more generally, time-to-event) functions.

Here are a few very recent examples. Judge, D.P. et al. (2025) use KMCs (as well as hazard functions within the context of a stratified Cox proportional hazards model) to study the efficacy of Acoramidis (compared to placebo) on all-cause mortality and cardiovascular-related hospitalization for patients with Transthyretin Amyloid Cardiomyopathy. They showed improvements in ACM and first CVH outcomes compared to the placebo. They report a divergence of the KMCs after about three months (but, possibly, a crossing of KMCs at the earlier stage was not mentioned). Ahn et al. (2025) compare the application of KMCs to the use of a Model for End-Stage Liver Disease numerical scale in a study of transplant-free survival. Li et al. (2025) used KMCs to compare 28-day mortality in individuals with sepsis in high and low hemoglobin, albumin, lymphocyte, and platelet (HALP) score groups. They report that a high HALP score decreases the 28-day, 90-day, and 360-day mortality (as well as in-hospital mortality) for patients with sepsis. However, they do not mention or discuss the problem of crossing KMCs that seems to occur in their KMCs at the early stages. Dai J. et al. (2025) used KMCs to compare overall and progression-free survival probabilities for stages B and C patients with intermediate to advanced hepatocellular carcinoma (HCC) who underwent transarterial chemoembolization (TACE) alone with patients who underwent a combination of TACE and ablation therapy. They looked at overall and progression-free survival. They concluded that for stages B and C patients, the overall survival probabilities were higher for those with combination therapy compared to those with only TACE therapy. The same was true for progression-free survival in stage B patients with the combination therapy, but this was not the case for stage C patients. Finally, in Yang, L. et al. (2025), Kaplan-Meier curves were employed to compare survival probabilities between patients with high and low TAP2 expression levels. Comparing KMCs for the two

expression levels allows for better patient prognostic assessment.[1]

A common issue that arises when comparing complementary cumulative distribution functions is that they often cross. But, in medical research, where these curves represent, for example, different treatments' survival functions, currently, the standard view is that crossing curves cannot be ranked. Thus, generally, non-crossing complementary cumulative distribution functions are a special case. It is, therefore, not surprising that researchers often find crossing KMCs in medical studies. For example, Kristiansen (2012) reviewed all publications in five major journals and found 175, about 47%, survival studies in which KMCs crossed (for example, crossing KMCs were found in Dehbi et al. (2018), Huang and Chong (2022), Tikidzhieva and Gentile and Magarotto et al. (2012)). Faced with this prevalent crossing KMCs problem, it became clear that this issue must be addressed. However, the medical literature has thus far not provided a solution to the problem of crossing KMCs. Indeed, Bouliotis and Billingham (2011) state, "There is a need in the clinical community to clarify methods that are appropriate when survival curves cross." They suggest that standard methods may not capture all aspects of the underlying survival function, so something else might be missing when KMCs cross. For example, they speculate that crossing curves might be related to differences in variances of the underlying distributions.

Some studies attempted to solve the crossing-curves problem by focusing not only on statistical issues or the heights of the KMCs but also on other characteristics of the KMCs. For example, Lin and Xu (2010) suggested a new survival function comparison method based on the absolute difference of the area under the KMCs over the whole range of the curves (namely, from the initial point to the endpoint). Similarly, Dormuth, Liu, et al. (2022) suggest that total areas (not the absolute differences) under the KMCs over the whole range (again, from the initial point to the endpoint, usually taken as the truncation point) should be used. Unfortunately (as shown in this paper), neither method is capable of comparing crossing survival distributions. The reason is that the comparison has to be "global," not "local." By local, it means areas up to a particular (single) point in time along the range of the distribution (in Lin Xu (2010)

[1] Recently, other methods of producing KMCs have been proposed. For example, Mamun et al. (2023) generate KMCs using machine learning methods and then use them for survival prediction in heart failure patients. Locey et al. (2024) provide the Rush Regression Workbench, which provides and automates a variety of commonly used statistical tools in medical research often used, including KMCs (and hazard functions). Nath et al. (2023) use an interactive web-based tool to obtain survival probabilities in brain cancer patients. Finally, Dag et al. (2023) also provide a machine learning approach, which they refer to as holistic. Specifically, they show that the prediction of lung cancer-related variables can be used to better explain lung cancer patients' survivability rates.

and Dormuth, Liu et al. (2022), this point is taken as the endpoint), whereas, global means the areas up to each point along the range of the distribution. Thus, for example, just because distribution A's area or absolute difference of the area from the initial point to the endpoint is higher than distribution B's, it does not follow that distribution A is better. In fact, as it is shown in the paper, comparing total areas under the KMCs from the initial to the endpoints is equivalent to comparing the means of the survival distributions, hence ignoring the differences in variances and all higher moments.

The main problem with studies applying the KMCs is that they are not cast within a decision-making framework. However, to choose a treatment is to make a decision, and to make decisions, one needs an objective function. By not casting the problem within a decision-making framework, insights that decision theory can provide are lost and left unexploited. These insights allow us to compare survival curves even if KMCs cross once or more.

The purpose of this paper is to cast the comparison of survival curves within a standard decision-making framework and use that framework to address the crossing KMCs problem. As an example, we apply the standard theory of decision under uncertainty, first developed by von Neumann and Morgenstern (von Neumann and Morgenstern, (1944)). Within this framework, common risk preference attributes and high-order stochastic dominance (SD) tools are used to compare SFs (for a good survey and discussion of stochastic dominance, see Levy (1992)). Specifically, higher-order SD tools enable us to extract information from KMCs that is otherwise left unexploited. Risk preference attributes allow us to use this extracted information to compare KMCs that are otherwise non-comparable. Such an approach is well-known in finance and economics and can provide the theoretical foundation for comparing KMCs. Consequently, in many cases, the crossing KMCs problem can be resolved. An intuitive way to view this approach is that it "substitutes" risk preference attributes for properties of the underlying SFs. Namely, as additional risk preference attributes are adopted, the restrictions required to rank survival functions conclusively become weaker (so no intersections may no longer be required). For example, when using "standard" KMCs (and given that - as is obvious - a later event occurrence - say death - is better), the requirement that one survival function should be everywhere higher than another is a necessary and sufficient condition for ranking the two survival functions. However, assuming the existence of risk aversion makes the necessary and sufficient condition for ranking two survival functions weaker, so they can be ranked even when

they cross.

Of course, all decision models require and are based on objective functions, and each objective function, in turn, is defined by its properties. The paper focuses on risk preference attributes because health-related decisions involve a great deal of uncertainty. The uncertainty is due to diagnostic and treatment risks, co-morbidity risks, potential future morbidity risks and background economics, financial and other risks. It is, therefore, clear that risk preference attributes play an important role in health-related decisions, such as treatment choices, as captured by comparing KMCs. Hence, as Attema, l'Haridon et al. (2019) state, "it is widely recognized that risk preferences are an important determinant of health-related decisions."[2]

This paper explores three possible risk preference attributes: risk aversion or discounting, prudence and temperance (in addition to the trivial property that a longer survival time is better than a shorter one). These properties are discussed in the next section. All three properties have been well-studied and demonstrated in numerous empirical studies in medical and other fields (e.g. finance, economics, political science, psychology and biology). The importance of risk aversion in health-related decisions has long been recognized in medical research (see, for example, Cykert, S., (2004), Klein and Stefanek (2007), Deakin, Alexander et al., (2009) and Riddel and Hales (2018)). Prudence was shown to be a primary determinant of the optimal level of prevention for health risks (see Courbage and Rey (2006), Felder and Mayrhofer (2014, 2017) and Eeckhoudt and Gollier (2005)). It was also shown that optimal medical test and treatment decisions are affected by prudence and temperance (see Felder and Mayrhofer (2014)). Consequently, Pauker (2014) recommended that these attributes be considered in medical research and decisions.

The paper shows that by using "high-order" risk preference attributes, well-known high-order SD tools can be applied to compare "standard" KMCs even if they intersect.[3] Specifically, a procedure that successively calculates "higher-order" KMCs whenever KMCs intersect is proposed. This procedure is based on higher-order stochastic dominance, which, in turn, is applicable if higher-order risk preference attributes are used.

These higher-order KMCs are obtained by calculating the areas under "lower order KMCs" up to each point in time over the whole range of the distribution. Thus, for example, the height

[2]See Attemaa, l'Haridonb et al. (2019).

[3]Using stochastic dominance in medical decisions was first mentioned by Leshno and Levy (2004). However, they did not provide real-life examples where SD (of any order) allows us to compare treatments in practice. Moreover, their methodological discussion was limited to first and second-order stochastic dominance.

of a KMC of order $n = 2, 3, 4$, at any point in time, represents the area under the KMC of order $n-1$ up to that point. Starting with "standard" KMCs, if, at any round, there is no intersection, there is no need to proceed, and the heights of KMCs (of that round) can be compared using existing methods.

On the other hand, if an intersection is encountered, the next-order KMCs are constructed, and their heights are compared. These steps are repeated until no intersection is found or stochastic dominance of order four is reached. If, after reaching fourth-order stochastic dominance, intersections are still found, we conclude that (given available risk preference attributes and hence admissible orders of stochastic dominance) it is impossible to rank the two treatments. Stochastic dominance of orders higher than four cannot be used without further knowledge of even higher order risk preference attributes (beyond temperance). Unfortunately, there is neither clear intuition regarding their meaning nor literature demonstrating their importance.

At each of these steps, the move to higher-order KMCs is facilitated by adopting more risk preference attributes, which allows us to apply stochastic dominance of a higher order. The paper provides necessary and sufficient conditions for ranking the high-order KMCs at each step.

Two examples based on empirical studies reported in the medical literature are provided to demonstrate this procedure. The first example is based on Nichols A. C. et al. (2022). In this example, first-order stochastic dominance is not satisfied, but second-order stochastic dominance is satisfied. For this example, an application of our method shows that even though the KMCs intersect, the preferred treatment can be determined if second-order stochastic dominance and risk aversion are applied. In the second example, based on Arenal, A. et al. (2022), both first-order and second-order stochastic dominance are not satisfied, but third-order stochastic dominance is satisfied. In this case, it is shown that even though the first and second-order KMCs intersect, the preferred treatment can be determined by applying third-order stochastic dominance, risk aversion and prudence.

2 Expected Utility and the Choice of Treatments

Consider treatments A and B. The treatments start at the same point in time, defined as zero. Under both treatments, an event may occur after the treatment begins. The event could be a biochemical recurrence, a clinical recurrence, death, etc. The time at which the event occurs is

denoted by x. Assume that x is a continuous random variable distributed over the closed interval $[0, b]$, whose probability distribution depends on the treatment given to a patient.[4] Define the density functions corresponding to treatments A and B as $f(x)$ and $g(x)$, respectively. For any point in time, z, define the survival functions[5] under treatments A and B as $H_1(z)$ and $J_1(z)$, respectively. To avoid confusion, note that subscript 1 reflects the fact that these SFs represent the areas under the density functions; hence, they are viewed as "first-order areas." The survival functions give us the probability that the event will occur *after* time z.

In their seminal contribution, Kaplan and Meier (1958) provided what became known as the Kaplan-Meier estimator or the KMC. Essentially, Kaplan and Meier used discrete data to estimate survival functions like $H_1(z)$ and $J_1(z)$ non-parametrically. Therefore, by their nature, the KMCs are discrete approximations of $H_1(z)$ and $J_1(z)$.

Given two survival functions,[6] to choose between treatments A and B, what is required for any decision problem is the decision maker's objective function. In our case, the decision maker is the patient.[7] The most obvious and commonly used framework that provides an objective function for choice under uncertainty is the expected utility model introduced by von Neumann and Morgenstern (1944)). While some reservations have been raised over the years about using this model (and several non-expected utility alternative models have been introduced), it has been and still is the main workhorse for dealing with choice under uncertainty, theoretically and empirically, including a wide range of medical fields. This paper also applies the expected utility model.

Thus, assume that the patient chooses a treatment by comparing the expected utilities of the outcomes, given the two distribution functions. Define the utility of outcomes (x) as $u(x)$.[8]

[4] The sample space of the two underlying distributions may not be the same, but following the stochastic dominance literature, we assume it is. So, for both treatments, the random variables are supported on the finite interval$[0, b],\ 0 < b < +\infty$. For a sufficiently large b, any reasonable possible upper bound for the two treatments can be accommodated.

[5] The corresponding cumulative distribution functions are $F(z) = \int_0^z f(x)dx$, $0 \leq z \leq b$ and $G(z) = \int_0^z g(x)dx$, $0 \leq z \leq b$. These CDFs give us the probability that the event will occur *before* time z. Define the corresponding complementary cumulative distribution functions (CCDFs) as $H_1(z) = 1 - F(z)$ and $J_1(z) = 1 - G(z)$. These are the probabilities that the event will occur after time z. The functions $H(z)_1$ and $J(z)_1$ are what we refer to as the survival functions. Finally, note that the means of x given the two distributions are given by $E_f(x) = \int_0^b xf(x)dx$ and $E_g(x) = \int_0^b xg(x)dx$.

[6] The survival curves themselves are determined by the underlying distribution functions, which are independent of the patient's preferences. It is the ranking of survival curves that depends on the patient's preferences. Therefore, for example, risk-averse and risk-neutral patients will rank the same survival curves, hence treatments, differently.

[7] The question of whether it should be the patient's or doctor's objective function is discussed, for example, in Makins (2023).

[8] Where $u(x)$ is continuous and differentiable of, at least, the fourth order.

Then, define the expected utilities of x under treatments A and B as $E_A[u(x)]$ and $E_B[u(x)]$, respectively. Since, in general, the survival functions of the two treatments differ, so do the corresponding expected utilities. The patient, therefore, chooses a treatment by comparing the expected utilities $E_A[u(x)]$ and $E_B[u(x)]$. The treatments' expected utilities can be written as $E_A[u(x)] = \int_0^b u(x)\, d[1 - H_1(z)]$ and $E_B[u(x)] = \int_0^b u(x)\, d[1 - J_1(z)]$.[9] Thus, a treatment's expected utility depends not only on the properties of its survival function but also on the properties of the patient's utility function. However, as is clear from the expected utilities' definitions, the properties of a treatment's survival function and the properties of the patient's utility function do not depend on each other.

Four common utility function properties have been observed and discussed in the theoretical and empirical decision theory literature, as well as medical research. This paper will refer to them as properties of orders one to four: P1-P4. Each of these properties implies a specific preference characteristic and thus gives rise to a particular type of behaviour when making decisions involving risk. The four properties are captured by the first four derivatives of the utility function (higher-order derivatives have not been considered in the literature since it is unclear what they represent). Denote these four derivatives as $u'(x)$, $u''(x)$, $u'''(x)$ and $u''''(x)$. The signs of these derivatives capture and imply a specific characteristic of an individual's preferences. Specifically, the signs of the second, third and fourth derivatives capture risk preference attributes (the second-order derivative may also capture future-discounting attributes). These four properties are now summarized below.

The first property (P1) is trivial; it says u must be an increasing function ($u'(x) > 0$). In other words, a later event occurrence is always preferred to an earlier one. This property simply says that "more time before the event occurs is better." If this property holds, the utility function is said to satisfy the first-order property, P1. The second property (P2) is risk aversion ($u''(x) < 0$). Risk aversion (with respect to the timing of the event occurrence) means that patients attach value to certainty and are willing to "pay a risk-premium" to avoid risk.[10]

[9] $E_A[u(x)] = \int_0^b u(x)\, f(x)dx = \int_0^b u(x)\, dF(x) = \int_0^b u(x)\, d[1 - H_1(z)]$ and $E_B[u(x)] = \int_0^b u(x)\, g(x)dx = \int_0^b u(x)\, dG(x) = \int_0^b u(x)\, d[1 - J_1(z)]$.

[10] An alternative interpretation of this property (which also implies that $u''(x) < 0$) is observationally equivalent) is that patients apply a discount factor to later years. The alternative interpretation of the second-order property is also intuitive and is related to "time-substitution" or "time discounting," in this case, current years versus future years. Generally, a year today is worth more than a year in the future. In the context of survival, this idea becomes even more important since people often think in terms of "quality years," and quality deteriorates over time; hence, later years are worth less than earlier ones. The discounting of future "rewards" underlies most economic and finance theories and appears even in neuroscience and animal behaviour (see Glasziou, Simes

Practically all of finance theory is based on the assumption that individuals are risk averse, and so are most economic models that deal with choice under uncertainty. The medical literature also recognizes the validity and importance of risk aversion (see, for example, Cykert, S., (2004), Klein and Stefanek (2007), Deakin, Alexander et al., (2009) and Riddel and Hales (2018)). It should be noted that all that matters is whether or not the patient is risk-averse, not the degree of risk aversion.

It has been recognized that individuals tend to exhibit prudence. Prudence refers to the sensitivity of an individual's actions in the presence of risk, whereas risk aversion refers to an individual's dislike of risk. Thus, prudence implies that an individual will take precautionary or preventive actions in the face of uncertainty. Prudence captures a "propensity to prepare and forearm oneself when facing uncertainty" (Kimball (1990)). The economic, finance and medical literature has noted and discussed its importance (Ebert and Wiesen (2011) and Crawford (2022)).

A classic example is that intertemporal saving decisions suggest that an individual will have precautionary savings. In the context of health, it means that an individual will act to ameliorate health risks. If prudence is satisfied, the utility function is said to satisfy the third-order property, P3 (prudence is satisfied if $u'''(x) > 0$).

Finally, temperance (P4) is another property of utility function that has been noted, addressed and used in economics and finance (for a finance example, see Colasante and Riccetti (2020); for an economics example, see Deck and Schlesinger (2010)). Temperance is satisfied if $u''''(x) < 0$. It means that an individual will reduce exposure to risky treatments if her independent background risk (from other risky choices) increases the degree of moderation in accepting risk. Several studies in various medical fields found that temperance plays an important role in patients' decisions regarding risky treatments while facing unrelated risks (see Felder and Mayrhofer (2014)).

The sections below show that by using "high-order" risk preference attributes P2-P4, well-known high-order SD tools can be applied to compare "standard" KMCs even if they intersect. As is clear from the definition of the expected utility functions, the more utility function proper-

and Gelber, (1990), Berridge and Kringelbach, (2015) and Schultz (2015)). For example, it was even suggested that the neurotransmitter serotonin might play an important role in modulating future discounting (Mobini, Chiang, et al. (2000)). There is also extensive literature that tries to untangle the contribution of risk aversion from the contribution of time substitution. We do not pursue this issue further in the paper. However, for our purpose, it suffices to note that both can explain P2.

ties we use, the less restrictive the survival functions can be. Alternatively, the more restrictive the survival functions, the fewer the number of utility function properties we need to use. In the standard Kaplan-Meier literature, no use is made of the utility function's properties P2-P4. Hence, the KMCs must indeed satisfy stringent conditions, i.e., no intersections are allowed.

3 First-order Stochastic Dominance and the Kaplan-Meier Curve

To place the KMC within the framework of choice under uncertainty, we start with the first-order stochastic dominance (SD1) concept. SD1 is defined as follows:

Definition 1: First-order Stochastic Dominance Treatment B exhibits SD1 over treatment A *if and only if* $J_1(z)$ is higher than $H_1(z)$ *for all* z values, *with strict inequality for at least some* z *values.*

Given Definition 1, the standard result is as follows:

Result 1: *If later event times are preferred (P1 holds), then treatment B is preferred to treatment A if and only* Treatment B exhibits SD1 over treatment A.

A discussion, derivation, and proofs of Result 1 and Results 2-4 below are given by Rothschild and Stiglitz (1970) and Levy (1992).

Figures 1a and 1b below are examples of survival functions where SD1 is satisfied and not satisfied, respectively.

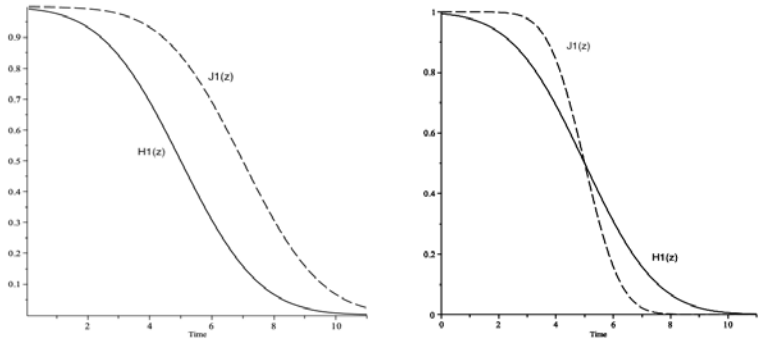

Figure 1a: SD1 is Satisfied. Figure 1b: SD1 is not satisfied

It is useful to note that if treatment B exhibits SD1 over treatment A, then it follows that the expected value of x under treatment B is necessarily higher than the expected value of x under treatment A.

Result 1 is appealing because other than assuming that "a later event date is preferred," no knowledge of other aspects of utility functions is required; all that is required is information about the SFs $J_1(z)$ and $H_1(z)$ at each point. However, it is important to emphasize that no conclusions can be drawn if $J_1(z)$ is higher than $H_1(z)$ over some range but not for all z values. In other words, this property must be global. This result stands in contrast with the standard Cox regression hazard ratio. Whereas in the Cox regression model, the hazard ratio is constant (by construction) - thus independent of time - here, for SD1 to hold, $J_1(z)$ needs to be higher than $H_1(z)$ for all z. Moreover, it is clear that if SD1 holds, any "height average" of $J_1(z)$ values will always be higher than the height average of $H_1(z)$ values. Equivalently, if SD1 holds, any "average hazard ratio" (an average of $J_1(z)/H_1(z)$) will be greater than 1, even though the hazard ratio for each value of z is not constant.

In practice, of course, one never knows the true underlying SFs; one can only work with the estimated $KMCs$. Hence, defining the $KMCs$ for treatments A and B as $KMC_1^A(x)$ and $KMC_1^B(x)$, respectively, then Result 1 implies that when later event times are preferred, then treatment B is preferred if and only if $KMC_1^B(z)$ is always (i.e., for all values of z) higher than $KMC_1^A(z)$, with strict inequality for at least some z values.

As is evident from the discussion above, SD1 of treatment B only holds under a very restrictive condition. Namely, $KMC_1^B(z)$ must be higher than $KMC_1^A(z)$ *for all values* of z. This clear ranking does not hold if, for example, for some values of z, $KMC_1^B(z)$ is higher than $KMC_1^A(z)$, but for others, it is lower. Such cases can occur if the KMCs intersect once or more. Although intersecting Kaplan-Meier curves are commonly found in the medical literature, there is still no consensus on the best way to deal with intersecting KMCs. In the following, a procedure is provided that addresses this problem by using higher-order properties of the utility function (properties P2-P4) and higher-order stochastic dominance concepts.

4 Second-Order KMCs

This section demonstrates that by using a further, higher-order property of the utility function, the severe restriction of SD1 can be relaxed, and the intersecting KMCs problem can then be

solved. Specifically, in addition to the obvious preference for a later event occurrence (P1), assume that utility functions are characterized by P2, namely, risk aversion. The restrictive SD1 - non-intersecting H and J curves - condition can now be relaxed with this additional assumption. It is now possible to conclude that, under certain conditions, treatment B is preferred to treatment A even if the J curve is not everywhere (for all z) higher than the H curve (hence, if $KMC_1^B(x)$ is not always above $KMC_1^A(x)$). In other words, even if the KMCs cross (once or more), treatment B may still be preferred to treatment A.

Define SD2 as follows:

Definition 2: Second-order Stochastic Dominance *Treatment B exhibits SD2 over treatment A if and only if the area under $J_1(z)$ is (weakly) greater than the area under $H_1(z)$ for all z (and strictly greater for at least some values of z).*

Given Definition 2, the standard result is as follows:

Result 2: *If a later event occurrence is better and, in addition, the utility function exhibits risk-aversion (i.e., both P1 and P2 hold), treatment B is preferred to treatment A if and only if Treatment B* exhibits SD2 over treatment *A.*

As Result 2 states, no local conclusions can be drawn, only global ones - the inequality in Result 2 must hold for every z over the whole range of z. Figures 2a and 2b show examples where SD2 is and is not satisfied, respectively. In Figure 2a, the area between the two curves before the intersection is greater than the area between the two curves after the intersection. Therefore, Result 2 implies that SD2 is satisfied. The opposite is true in Figure 2b.

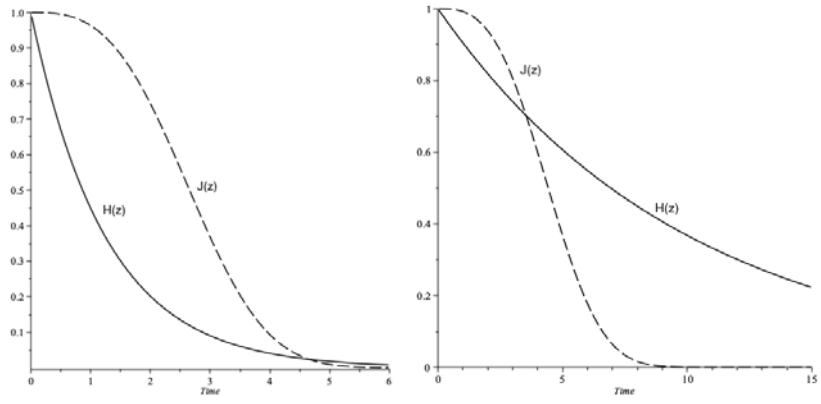

Figure 2a: SD2 is Satisfied Figure 2b: SD2 is Not Satisfied

Result 2 provides the basis for comparing treatments by calculating the areas under the Kaplan-Meier Curves when those curves intersect (once or more). Hence, Result 2 tells us that

treatment B is preferred to treatment A if and only if the area under $KMC_1^B(z)$ is (weakly) greater than the area under $KMC_1^A(z)$ for all values of z (and, strictly greater for at least some z values). This result holds even if there are intersections.

The idea of calculating the areas under KMCs to compare treatments was introduced in medical studies before (for example, see Lin and Xu (2010), Dormuth, Liu, et al. (2022)). However, in these "restricted mean survival time" (RMST) studies, the areas under KMC^B and KMC^A were calculated only for one specific value of z; the endpoint (usually, at the truncation point), but not for every z, as required by Result 2. So, for example, if the truncation was at point t, the RMST studies compared the areas under the $KMC^B(t)$ and $KMC^A(t)$ curves up to point t, but not for all values. Therefore, Result 2 implies that it cannot be determined which treatment is better. What does such a comparison tell us? Since the difference in areas under cumulative distributions over the entire range is equal to the difference in means (see Rothschild and Stiglitz (1970)), the RMST studies simply tell us which distribution has a higher mean. The fact that such a comparison is insufficient to compare treatments is not surprising since a comparison of means ignores all higher moments of the distributions. For example, suppose the means of the two distributions are the same. Then, it can be shown that the difference in the total areas under the cumulative distributions (over the whole range) is simply the difference in the variances.[11] Therefore, it is clear that a (local) comparison of the areas under the $KMC_1^B(z)$ and $KMC_1^A(z)$ curves up to point z but not for all values below z ignores information on higher orders moments, e.g., second moments.

The following procedure is applied to make Result 2 visually more intuitive, simpler to apply, and easier to relate to standard KMCs. First, define the areas under $H_1(z)$ and $J_1(z)$, up to the point z, as $H_2(z)$ and $J_2(z)$. In other words, the heights of $H_2(z)$ and $J_2(z)$ measure the areas under the $H_1(z)$ and $J_1(z)$ curves (which themselves are areas under the density functions - hence the subscript 2)[12]. Second, define the difference between the hights of $J_2(z)$ and $H_2(z)$ as $D_2(z) = J_2(z) - H_2(z)$. The $D_2(z)$ curve is referred to as the second-order difference curve. Result 2 can now be stated in terms of the single difference curve $D_2(z)$. Specifically,

Result 2a: *If a later event occurrence is better and, in addition, the utility function exhibits*

[11] If we define the variances of x under the f and g disdributions as $Var_F(x)$ and $Var_G(x)$, then (using integration by parts), $Var_F(x) - Var_G(x) =$
$2\int_0^b[\int_0^z\{F(z) - G(z)\}dz]dz$.
Therefore, SD2 of B over A implies that $Var_F(x) \geq Var_G(x)$. However, the converse is not true.

[12] $H_2(z) \equiv \int_0^z H(x)\,dx$ and $J_2(z) \equiv \int_0^z J(x)\,dx$.

risk-aversion (i.e., both P1 and P2 hold), treatment B is preferred to treatment A if and only if $D_2(z)$ is higher than zero for all z (and strictly greater than zero for at least some values of z).[13]

Result 2a allows us to compare treatments by simply comparing the height of the single $D_2(z)$ curve for each z rather than comparing the areas under the $H_1(z)$ and $J_1(z)$ curves up to each z. Therefore, to compare treatments whose Kaplan-Meier Curves intersect, first define the areas under $KMC_1^A(z)$ and $KMC_1^B(z)$ up to point z as $KMC_2^A(z)$ and $KMC_2^B(z)$, respectively. The heights of $KMC_2^A(z)$ and $KMC_2^B(z)$ at a particular z now represent the area under $KMC_1^A(z)$ and $KMC_1^B(z)$ up to that z. $KMC_2^A(z)$ and $KMC_2^B(z)$ can be viewed as "*second-order KMCs*." Next, for each z, define the difference between the heights of $KMC_2^B(z)$ and $KMC_2^A(z)$ as $D_2^{KMC}(z) = KMC_2^B(z) - KMC_2^A(z)$. This curve is referred to as the second-order Kaplan-Meier Difference Curve.

Result 2a can now be restated in terms of the height of the $D_2^{KMC}(z)$ curve as follows:

Result 2b: *If a later event occurrence is better and, in addition, the utility function exhibits risk-aversion (i.e., both P1 and P2 hold), treatment B is preferred to treatment A if and only if $D_2^{KMC}(z)$ is greater than zero for all z (and, strictly higher than zero for at least some values of z).*

As was the case above, local comparisons cannot be made, only global ones.

Since estimation methods and tests for the significance of height differences for $KMC_1^B(z)$ and $KMC_1^A(z)$ already exist, presumably, similar estimation techniques and tests can also be used for second-order KMCs or the $D_2^{KMC}(z)$ curve. The use of second-order KMCs or the corresponding $D_2^{KMC}(z)$ curve addresses the crossing curves problem. It allows us to compare treatments by exploiting the fact that (in addition to P1) the utility function satisfies $P2$.

From Definitions 1 and 2, it is clear that SD1 imposes stronger restrictions than SD2. Specifically, SD1 implies SD2, but the reverse is not true. Hence, Result 2 (or its variants) allows us to compare treatments A and B under weaker restrictions; namely, it can be used even if Result 1 cannot be applied.

Finally, two points are worth noting. First, if the patient is not risk-averse, we cannot apply Results 2, 2a and 2b. Hence, we are back to the standard Kaplan-Meier problem where crossing

[13] Note that when $z = b$, the condition $J_2(b) \geq H_2(b)$, whish is the same as the condition $D_2(b) \geq 0$, implies that $E_g(x) \geq E_f(x)$.

KMCs cannot be ranked. Second, all that matters is whether risk aversion exists, not its level.

5 Third-order KMCs

Now, what happens if the $KMC_2^B(z)$ and $KMC_2^A(z)$ intersect? Namely, if $D_2^{KMC}(z)$ is not always positive. If intersections occur, the method discussed above can be used to relax further the conditions the survival functions must satisfy. This can be done by assuming that individuals also exhibit prudence (P3). With prudence, the requirement that $J_2(z)$ must be higher than $H_2(z)$ for all z can be weakened, thus allowing for their intersections.

Define third-order stochastic dominance (SD3) as follows:

Definition 3: Third-order Stochastic Dominance *Treatment B exhibits SD3 over treatment A if and only if the area under $J_2(x)$ is greater than the area under $H_2(x)$ for all z* values *(and strictly greater for at least some z), and $D_2(b) = J_2(b) - H_2(b) \geq 0$.*[14]

Given Definition 3, the standard result is as follows:

Result 3: *If a later event occurrence is better and, in addition, the utility function exhibits risk-aversion and prudence (i.e., P1, P2 and P3 all hold), treatment B is preferred to treatment A, if and only if treatment B exhibits SD3 over treatment A.*

The distributional restrictions imposed by Result 3 are weaker than those in Result 2, which allows us to compare the treatments even when $J_2(z)$ and $H_2(z)$ intersect (when $D_2(z)$ is not always positive).

Again, to make the implication of Result 3 visually more intuitive and easier to apply to KMCs, differences-in-area-curves can be used. Hence, first, define the areas under $H_2(z)$ and $J_2(z)$ as $H_3(z)$ and $J_3(z)$, respectively.[15]. Next, define the difference between $J_3(z)$ and $H_3(z)$ as $D_3(z) = J_3(z) - H_3(z)$. Result 3 can now be stated in terms of the hight-difference curve, $D_3(z)$, rather than the areas under the $J_2(z)$ and $H_2(z)$ curves used in Result 3. Result 3 can then be restated as follows,

[14] Remember that the condition $D_2(b) \geq 0$ implies that $E_g(x) \geq E_f(x)$, i.e., the mean of x under treatment B is higher than the mean under treatment A.

[15] $H_3(z)$ and $J_3(z)$ measure the areas under $H_2(z)$ and $J_2(z)$, which, in turn, are areas under $H_1(z)$ and $J_1(z)$, but $H_1(z)$ and $J_1(z)$ are areas under the density functions. Hence the subscript 3.

Result 3a: *If a later event occurrence is better and, in addition, the utility function exhibits risk-aversion and prudence (i.e., P1, P2 and P3 all hold), treatment B is preferred to treatment A, if and only if* $D_3(z)$ is greater than zero *for all z (and strictly higher than zero for at least some z) and* $D_2(b) \geq 0$.

Therefore, to compare treatments whose second-order Kaplan-Meier Curves intersect, first define the areas under $KMC_2^A(z)$ and $KMC_2^B(z)$ as $KMC_3^A(z)$ and $KMC_3^B(z)$, respectively. $KMC_3^A(z)$ and $KMC_3^B(z)$ can be viewed as "*third-order KMCs*." The heights of $KMC_3^A(z)$ and $KMC_3^B(z)$ now represents the areas under $KMC_2^A(z)$ and $KMC_2^B(z)$, respectively. Next, define the difference of the heights of $KMC_3^B(z)$ and $KMC_3^A(z)$ as $D_3^{KMC}(z) = KMC_3^B(z) - KMC_3^A(z)$. Result 3a can now be stated as follows:

Result 3b: *If a later event occurrence is better and, in addition, the utility function exhibits risk-aversion and prudence (i.e., P1, P2 and P3 all hold), treatment B is preferred to treatment A, if and only if* $D_3^{KMC}(z)$ *is greater than zero for all z (and strictly greater for at least some z) and* $D_2^{KMC}(b) \geq 0$.

Definitions 3a and 2a clearly show that SD2 imposes stronger restrictions than SD3. Specifically, SD2 implies SD3, but the converse is not true. Hence, Result 3b allows us to compare treatments A and B under weaker restrictions than the ones in Result 2b; namely, it can be used even if Result 2b cannot be used (even if the second-order KMCs intersect). Moreover, all that matters is whether prudence exists, not its level. Now, what happens if the patient is not prudent? Unfortunately, in such a case, we cannot apply Results 3, 3a and 3b.

In principle, if SD3 is not satisfied (if $J_3(z)$ and $H_3(z)$ intersect), the restrictions imposed on the survival functions can be weakened even further if patients also exhibit temperance (P4). The adoption of the temperance assumption allows us to compare the two treatments by considering differences in the heights of "*fourth-order KMCs*" (defined similarly as the third-order KMCs defined above) or the positivity of the *fourth-order* difference curve.[16] Since we were unable to find real medical data for which $D_3^{KMC}(z)$ *is not greater than zero for all* z, but the corresponding fourth-order difference curve is, this case is not pursued here.

[16] The use of fourth-order stochastic dominance is described as follows. Define the areas under $KMC_3^B(z)$ and $KMC_3^A(z)$ as $KMC_4^B(z)$ and $KMC_4^A(z)$, and the corresponding difference as $D_4^{KMC}(z) = KMC_4^B(z) - KMC_4^A(z)$. Then, we have the following result:

Result 4b: If P1, P2, P3 and P4 all hold, then treatment B is preferred to treatment A if and only if $D_4^{KMC}(z)$ *is greater than zero for all z (and strictly higher for at least some z) and* $D_2^{KMC}(b) \geq 0$.

Note that in all cases, high-order difference curves allow us to check for the positivity of a single curve rather than compare areas under two curves. Consequently, existing statistical testing methods (available for standard - $KMC_1^B(z)$ and $KMC_1^A(z)$ - curves) can be used to compare high-order KMCs and check for the positivity of difference curves.

In summary, the following procedure is proposed. First, if $KMC_1^B(z)$ and $KMC_1^A(z)$ intersect, obtain $KMC_2^A(z)$ and $KMC_2^B(z)$ by measuring the area under $KMC_1^B(z)$ and $KMC_1^A(z)$ up to each point in time. These curves are used to define the second-order difference curve $D_2^{KMC}(z)$. Then, check for the positivity of the $D_2^{KMC}(z)$ curve. If $D_2^{KMC}(z)$ is sometimes smaller than zero, repeat the step above and obtain third-order KMCs by measuring the area under the second-order KMCs up to each point in time for all z. These are used to define the third-order difference curve $D_3^{KMC}(z)$. If $D_3^{KMC}(z)$ is always positive (and $D_2^{KMC}(b)$ is positive), the conclusion is that the corresponding treatment is preferred. If there still is an intersection, repeat the steps above to compare fourth-order KMCs or the fourth-order difference curve. If the fourth-order difference curve $D_4^{KMC}(z)$ is always positive (and $D_2^{KMC}(b)$ is positive), the conclusion is that treatment B is preferred. If $D_4^{KMC}(z)$ is not always positive, the conclusion is that ranking the two alternative treatments is impossible. As we go from first-order KMCs to higher-order ones, additional attributes of risk preferences are used. The extra information obtained from the underlying preference allows us to use higher-order SD to obtain all the extractable information embedded in the KMCs.

The following section provides examples of this procedure using actual medical treatment data.

6 Empirical Examples

This section applies our methodology to two examples of crossing KMCs in published medical papers. Following our procedure, first, the differences in the areas under the two treatments' KMCs (for all z values) are calculated. In other words, the differences between $KMC_2^B(z)$ and $KMC_2^A(z)$ up to point z for all z (defined above as $D_2^{KMC}(z)$) are calculated. If this difference is always positive, the conclusion is that SD2 is satisfied. Hence, if a later event occurrence is preferred and the utility function exhibits risk aversion (when Properties P1 and P2 are satisfied), treatment B is preferred to treatment A. This case occurs in Example 1 below and is shown in Figures 3a and 3b.

If SD2 is not satisfied, the differences between $KMC_3^B(z)$ and $KMC_3^A(z)$ up to point z for all z (defined above as $D_3(z)$) are calculated. Again, if this difference is always positive for all z (and $D_2^{KMC}(b)$ is positive), the conclusion is that SD3 is satisfied. Thus, *if a later event occurrence is better and, in addition, the utility function exhibits risk aversion and prudence* (when Properties P1, P2 and P3 are satisfied), treatment B is preferred to treatment A. This case occurs in Example 2 and is shown in Figures 4a, 4b and 4c below.

As the examples below demonstrate, even if the two KM curves intersect, it is still valuable to check for higher-order stochastic dominance, which provides additional information about the compared treatments. The additional information may allow us to discriminate between two treatments that are not comparable when looking only at first-order KMCs.

6.1 Example 1: SD1 is Not Satisfied, SD2 is Satisfied

The first example is based on Nichols A. C. et al. (2022). The study compares the long-run outcomes of radiotherapy (RT - with chemotherapy if N1-2) with those of transoral robotic surgery plus neck dissection (TRND),[17] in the treatment of oropharyngeal squamous cell carcinoma.[18]

In the following, TRND is defined as treatment A and RT as treatment B. The corresponding KMCs are shown in Figure 3a. As this figure shows, the KMCs cross so that SD1 is not satisfied. Hence, following our procedure, first, the areas under the KMCs as defined by $KMC_2^B(z)$ and $KMC_2^A(z)$ up to point z for all z. These areas were then used to calculate the differences between them as $D_2^{KMC} = KMC_2^B(z) - KMC_2^A(z)$. Figures 3a and 3b show that while SD1 is not satisfied, SD2 is.

[17] With or without adjuvant therapy.

[18] The Method: Authors randomly assigned patients with T1-T2, N0-2 ($\leq$ 4 cm) OPSCC to radiotherapy (RT) (with chemotherapy if N1-2) versus transoral robotic surgery plus neck dissection (TORS + ND) (with or without adjuvant therapy). The primary endpoint was swallowing quality of life (QOL) at 1 year using the MD Anderson Dysphagia Inventory. Secondary endpoints included adverse events, other QOL outcomes, overall survival, and progression-free survival. All analyses were intention-to-treat.

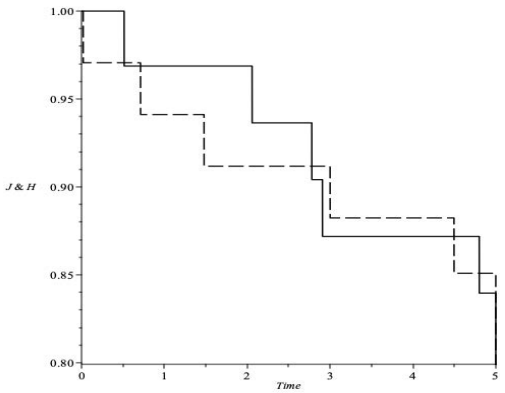

Figure 3a. KMCs. RT- Solid Line, TRND - Dashed Line.

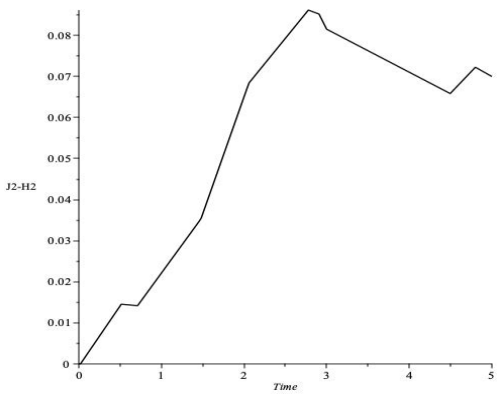

Figure 3b: D_2^{KMC}.

This result is demonstrated in Figure 3a since the area between the solid and dashed lines where the solid line is higher than the dashed line dominates the area between the dashed and solid line where the dashed line is higher than the solid line. It is also demonstrated in Figure 3b, where the difference is positive ($D_2^{KMC} > 0$) for all z. Hence, whereas the treatments could not be ranked by looking at SD1, they can be ranked by using SD2 and assuming that (in addition to P1) risk aversion is satisfied. In other words, the conclusion is that the RT treatment is preferred to the TRND treatment.

6.2 Example 2: SD1 and SD2 Are Not Satisfied, SD3 is Satisfied

The third example is based on Arenal, A. et al. (2022). This study compares the efficacy and safety of catheter ablation (CA) and antiarrhythmic drugs (AADs) as first-line therapy in patients with ischemic cardiomyopathy and an implantable cardioverter-defibrillator (ICD). Treatment ICD is defined as treatment A and CA as treatment B.

The corresponding KMCs are shown in Figure 4a. Again, as this figure shows, the KMCs cross so that SD1 is not satisfied. Hence, following our procedure, the areas under the KMCs (as defined by $KMC_2^B(z)$ and $KMC_2^A(z)$ up to point z for all z) are calculated. Then, the differences between them, defined above as D_2^{KMC}, are calculated. Figure 4b shows the differences between $KMC_2^B(z)$ and $KMC_2^A(z)$ up to point z for all z. As the figure shows, SD2 is also not satisfied (but $D_2^{KMC}(b)$ is positive). Thus, the areas under $KMC_2^B(z)$ and $KMC_2^A(z)$, are calculated which gives us the $KMC_3^B(z)$ and $KMC_3^A(z)$ curves. The difference between $KMC_3^B(z)$ and $KMC_3^A(z)$ curves, defined as D_3^{KMC}, is shown in Figure 4c. As this figure shows, $D_3^{KMC} > 0$ for all z (and Figure 4b shows that $D_2^{KMC}(b)$ is positive). Hence, whereas the treatments could not be ranked by using SD1 or SD2, they can be ranked by using SD3 and

assuming that (in addition to P1) risk aversion and prudence are satisfied. In other words, the conclusion is that the CA treatment (treatment B) is preferred to the ICD treatment (treatment A).

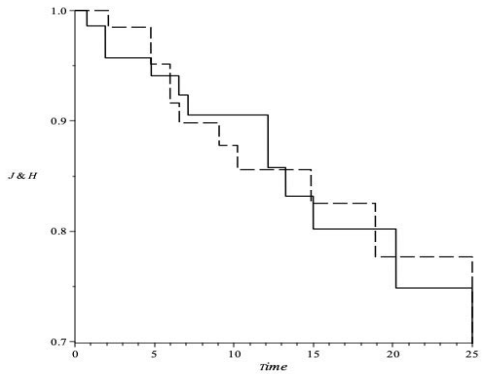

Figure 4a: CA & ICD, Dashed and Solid Lines

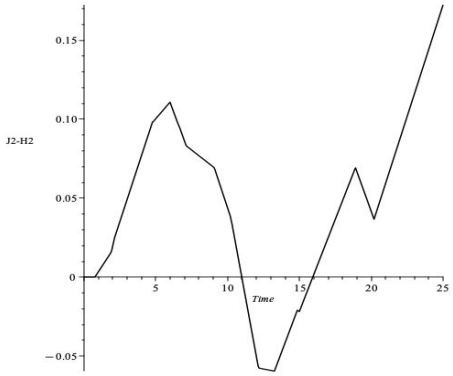

Figure 4b: $D_2^{KMC}(z)$.

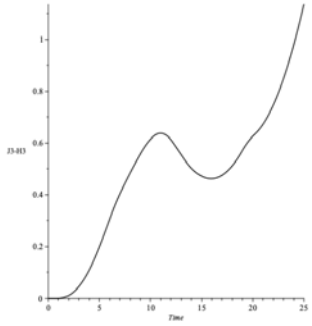

Figure 4c: $D_3^{KMC}(z)$.

7 Critical Discussion and Limitations

KMCs have been used extensively in countless applications in various medical fields to compare survival functions. It has been recognized in the medical literature that KMCs frequently cross. However, currently, no framework exists in the medical literature that allows us to use KMCs when they cross.

This paper provides a solution to the problem of crossing KMCs. It does so by adopting a standard decision-making framework. It is this framework that allows us to apply stochastic dominance concepts and risk preference attributes to solve the crossing KMCs problem. Furthermore, the solution this framework provides is simple, practical and easily applicable.

However, our approach has several shortcomings. First, like all theoretical models, the strength of our model depends on the validity of its underlying assumptions, specifically, the

applicability of the expected utility framework and the validity of the assumptions regarding risk preference attributes. While the expected utility assumption may and has been challenged, it is still the standard approach for problems of choice under uncertainty in a wide range of fields. Similarly, the assumptions regarding risk preference attributes are widely accepted and have been confirmed in empirical studies in many disciplines, including economics, finance, psychology, biology, and even medicine. Therefore, the theoretical and empirical underpinnings of our model are quite solid. Moreover, it is easy to elicit some risk preference information in a doctor-patient setting by presenting patients with a simple questionnaire (for example, regarding risk-aversion: do you prefer a treatment that gives you an extra 5 years for sure or one that gives you an extra 2 or 8 years with equal probabilities, (ii) regarding prudence: will you take precautionary or preventive actions in the face of medical risks?).

Second, our model cannot always guarantee a solution. The reason is that information on risk preference attributes of orders higher than four does not exist. Consequently, once SD4 is reached and an intersection still exists, it is impossible to go further. Third, if a specific risk preference attribute does not hold for a particular patient (e.g., say, prudence), the particular corresponding SD concept cannot be applied.

Finally, our paper does not provide explicit testing procedures. However, existing testing techniques could presumably be used. The challenging global testing question is beyond the scope of this paper and can be addressed in future research.

8 Conclusion

KMCs have been extensively used in the medical research. Yet, the problem of crossing KMCs has not been solved. Indeed, it is clear that without a theoretical framework, this problem is challenging, if not impossible, to solve. Thus, in the literature, whenever KMCs cross, a descriptive evaluation, rather than a ranking, is usually provided. This paper integrates the comparison of survival curves into decision theory, thus providing a theoretical framework for comparing crossing KMCs. Framing the survival function comparison as a decision-theoretic problem allows us to use risk preference attributes in conjunction with high-order stochastic dominance tools to solve the problem of crossing KMCs. This paper provides a procedure that uses successively higher-order KMCs in conjunction with higher-order risk preferences to compare treatments whenever KMCs intersect. This process is possible because it uses otherwise

unexploited higher-order properties of the underlying Kaplan-Meier Curves and utility functions. Since our knowledge of risk preference attributes is limited (it does not exceed the fourth order), our procedure cannot necessarily always guarantee a solution. We apply our methodology to two examples of crossing KMCs in published medical papers and show that, in these examples, treatments that otherwise could not be ranked can be ranked by our procedure.

8.1 References

Ahn, J.C. et al. (2025), "AI-driven synthetic data generation for accelerating hepatology research: A study of the United Network for Organ Sharing (UNOS) database," Hepatology, 10-1097.

Arenal, A., et al. (2022). "Substrate Ablation vs Antiarrhythmic Drug Therapy for Symptomatic Ventricular Tachycardia." J Am Coll Cardiol 79(15): 1441-1453.

Attema, A.E., l'Haridon, O., et al., (2019), "Measuring multivariate risk preferences in the health domain," Journal of Health Economics 64, 15-24.

Berridge, C, and M.L. Kringelbach, (2015), "Pleasure Systems in the Brain," Neuron, 86, 3, 646-664. DOI:https://doi.org/10.1016/j.neuron.2015.02.018

Bouliotis, G. and L, Billingham, (2011), "Crossing survival curves: alternatives to the log-rank test," Trials 12 (Suppl 1), A137. https://doi.org/10.1186/1745-6215-12-S1-A137.

Colasante, A. and L. Riccetti, (2020), "Risk aversion, prudence and temperance: It is a matter of gap between moments," Journal of Behavioral and Experimental Finance, 25, p.100262.

Courbage, C. and B. Rey, (2012), "Priority setting in health care and higher order degree change in risk," Journal of Health Economics, 31, 484-489.

Crawford, S., (2022), "Evaluating Treatment Options for Metastatic, Castration-Resistant Prostate Cancer: A Comprehensive Value Assessment," Doctoral dissertation, University of Southern California.

Cykert, S., (2004), "Risk acceptance and risk aversion: patients' perspectives on lung surgery," Thoracic surgery clinics, 14, 3, 287-293.

Dai, Jingyi, et al., (2025), "Survival Benefits of Transarterial Chemoembolization Plus Ablation Therapy in Patients With Intermediate or Advanced Hepatocellular Carcinoma: A Propensity Score Matching Study," Cancer Management and Research, 483-497.

Deakin, C.T., Alexander, I.E. and I. Kerridge, (2009), "Accepting Risk in Clinical Research:

Is the Gene Therapy Field Becoming Too Risk-averse?" Molecular Therapy, 17, 11, 842-1848, DOI:https://doi.org/10.1038/mt.2009.223.

Deck, C. and H. Schlesinger, (2010), "Exploring higher order risk effects," The Review of Economic Studies, 77, 4, 1403-1420.

Dehbi, H.M., Royston, P. and A. Hackshaw, (2018), "Life expectancy difference and life expectancy ratio: two measures of treatment effects in randomised trials with non-proportional hazards," BMJ, 2017;357:j2250, doi: https://doi.org/10.1136/bmj.j2250

Dormuth, I., Liu, T., Xu, J. et al., (2022), "Which test for crossing survival curves? A user's guideline." BMC Med Res Methodol 22, 34. https://doi.org/10.1186/s12874-022-01520-0.

Ebert, S., and D. Wiesen, (2011), "Testing for Prudence and Skewness Seeking," Management Science, 57, 7, 1334-1349.

Eeckhoudt, L., and C. Gollier, (2005), "The Impact of Prudence on Optimal Prevention, "Economic Theory, 26, 4, 989-994.

Felder, S. and T Mayrhofer, (2014), "Risk preferences: consequences for test and treatment thresholds and optimal cutoffs," Medical Decision Making 34, 33–41.

Gentile, M., Magarotto, V. et al. (2017), "Lenalidomide and low-dose dexamethasone (Rd) versus bortezomib, melphalan, prednisone (VMP) in elderly newly diagnosed multiple myeloma patients: a comparison of two prospective trials," American Journal of Hematology, 92, 3, pp. 24-250.

Glasziou, P. P., R.G. Simes and R. D. Gelber, (1990), "Quality adjusted survival analysis," Statistic in Medicine, 9, 11, 1259-1276.

Huang, Q. and C. Tian, (2022), "Visualizing Time-Varying Effect in Survival Analysis: 5 Complementary Plots to Kaplan-Meier Curve," Oxidative Medicine and Cellular Longevity, vol. 2022, Article ID 3934901, https://doi.org/10.1155/2022/3934901.

Judge, D.P., et al. (2025), "Efficacy of acoramidis on all-cause mortality and cardiovascular hospitalization in transthyretin amyloid cardiomyopathy." Journal of the American College of Cardiology, 85.10, 1003-1014.

Kaplan, E. L. and P. Meier, (1958), "Nonparametric estimation from incomplete observations," Journal of the American Statistical Association, 53, 282, 457-481.

Kimball, M. S., (1990), "Precautionary Saving in the Small and in the Large," Econometrica, 58, 53-73.

Klein, W. M., and M.E., Stefanek, (2007), "Cancer risk elicitation and communication: lessons from the psychology of risk perception," CA: a cancer journal for clinicians, 57, 3, 147-167.

Kristiansen, I.S., (2012), "PRM39 Survival curve convergences and crossing: a threat to the validity of meta-analysis?," Value in health, 15, 7. A652 10.1016/j.jval.2012.08.290.

Locey, K. et al. (2024), "Rush regression workbench: An integrated open-source application for regression modeling and analysis in healthcare analytics," Healthcare Analytics, 5, https://doi.org/10.1016/j.health.2024.100314

Leshno M. and H. Levy, (2004), "Stochastic dominance and medical decision making," Health Care Manag Sci. 2004 Aug;7(3):207-15. doi: 10.1023/b:hcms.0000039383.54736.65. PMID: 15648563.

Levy, H, (1992), "Stochastic Dominance and Expected Utility: Survey and Analysis," Management Science, 38, 4, 555-593.

Li, H. et al., (2025), "The relationship between hemoglobin, albumin, lymphocyte, and platelet (HALP) score and 28-day mortality in patients with sepsis: a retrospective analysis of the MIMIC-IV database." BMC Infectious Diseases 25,1, 333.

Lin, X, and Q., Xu, (2010), "A new method for the comparison of survival distributions," Pharm Stat 9, 1, 67–76. 10.1002/pst.376

Makins, N., (2023), "Patients, doctors and risk attitudes," Journal of Medical Ethics, 49, 11, 737-741.

Mamun, Md.A. et al. (2023), "A machine learning approach for risk factors analysis and survival prediction of heart failure patients," Healthcare Analytics, https://doi.org/10.1016/j.health.2023.100182

Mobini, S., Chiang, T.J., et al. (2000), "Effect of central 5-hydroxytryptamine depletion on inter-temporal choice: A quantitative analysis," Psychopharmacology, 149, 3, 313–318. doi:10.1007/s002130000385.

Nath, P.G. et al. (2023), "An interactive web-based tool for predicting and exploring brain cancer survivability," Healthcare Analytics, 3, https://doi.org/10.1016/j.health.2022.100132.

Nichols A. C. et al. (2022), "Randomized Trial of Radiotherapy Versus Transoral Robotic Surgery for Oropharyngeal Squamous Cell Carcinoma: Long-Term Results of the ORATOR Trial." Journal of Clinical Oncology, 40, 8, https://doi.org/10.1200/JCO.21.01961.

Pauker, S.G., (2014), "Moments When Utilities Are Functional," Medical Decision Making,

34, 4-7. https://doi.org/10.1177/0272989X13495714.θ

Riddel, M. and D. Hales, (2018), "Predicting Cancer-Prevention Behavior: Disentangling the Effects of Risk Aversion and Risk Perceptions," Risk Analysis, 38, 10.

Rothschild, M. and J. E. Stiglitz, (1970), "Increasing risk: I. A definition," Journal of Economic Theory, 2, 3, 225-243.

https://doi.org/10.1016/0022-0531(70)90038-4.

Schultz, W., (2015), "Neuronal Reward and Decision Signals: From Theories to Data," Physiological Reviews., 95, 3, 853–951. doi:10.1152/physrev.00023.2014.

Tikidzhieva, A., Benner, A., M. S., Formentini, et al., (2012), "Microsatellite instability and Beta2-Microglobulin mutations as prognostic markers in colon cancer: results of the FOGT-4 trial," British Journal of Cancer, 106, 6, 1239–1245. 10.1038/bjc.2012.53.

Von Neumann, J. and O. Morgenstern, (1944), Theory of Games and Economic Behavior, Princeton University Press.

Yang, L. et al., (2025) Identification of TAP2 as a novel immune target in human cancers: insights from integrated bioinformatics and experimental approaches. Eur J Med Res 30, 163 (2025). https://doi.org/10.1186/s40001-025-02360-6