Published in final edited form as:

Biometrics. 2018 March ; 74(1): 86–99. doi:10.1111/biom.12740.

A Local Agreement Pattern Measure Based on Hazard Functions for Survival Outcomes

Tian Dai¹, Ying Guo^{1,*}, Limin Peng¹, and Amita K. Manatunga¹

¹Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, Atlanta, Georgia 30322, U.S.A

Summary

Assessing agreement is often of interest in biomedical and clinical research when measurements are obtained on the same subjects by different raters or methods. Most classical agreement methods have been focused on global summary statistics, which cannot be used to describe various local agreement patterns. The objective of this work is to study the local agreement pattern between two continuous measurements subject to censoring. In this paper, we propose a new agreement measure based on bivariate hazard functions to characterize the local agreement pattern between two correlated survival outcomes. The proposed measure naturally accommodates censored observations, fully captures the dependence structure between bivariate survival times and provides detailed information on how the strength of agreement evolves over time. We develop a nonparametric estimation method for the proposed local agreement pattern measure and study theoretical properties including strong consistency and asymptotical normality. We then evaluate the performance of the estimator through simulation studies and illustrate the method using a prostate cancer data example.

Keywords

Agreement; Bivariate survival times; Hazard functions; Kernel smoothing

1. Introduction

In biomedical studies, researchers are often interested in assessing agreement on measurements taken on the same subjects using different methods or by different raters. There has been extensive literature on assessing agreement and making appropriate inference. For continuous data, various measures including both scaled measures, such as concordance correlation coefficient (CCC) (Lin et al., 2002) and unscaled ones, such as total deviation index (Lin et al., 2002), have been proposed. For categorical data, the kappa coefficient and its extensions (Cohen, 1960, 1968) have been widely used.

* yguo2@emory.edu.

7. Supplementary Materials

Web Appendices and Tables referenced in Sections 1, 2, 3 and 5 are available with this paper at the *Biometrics* website on Wiley Online Library.

All of the aforementioned methods take the strategy of quantifying the agreement of interest by a global summary measure. While being simple, they have been criticized for their limitations in fully capture agreement information (Tanner and Young, 1985; Darroch and McCloud, 1986). For example, with two categorical scales, Agresti (1989) showed that when a simple quasi-symmetry model holds for the contingency table, the kappa contains all relevant information about the structure of the agreement. However, when the quasisymmetry model fails, various agreement patterns can produce the same kappa value and the kappa coefficient alone is not capable of distinguishing different agreement patterns. Given the limitation of kappa coefficient, Tanner and Young (1985), Agresti (1988) and others have promoted studying the structure of agreement instead of summarizing agreement into a single measure. To this end, they proposed alternative approaches based on log-linear modeling. For continuous data, researchers (Borg et al, 1995; Schild et al., 2000) have argued to use descriptive tools such as a scatter plot of two continuous readings with the 45 degree line as the concordance line or Bland and Altman plot of the difference against the average of paired measurements with the horizontal line of zero as the reference. These plots can be informative, in particular in visualizing the local patterns of agreement, and help capture the changes of measurement concordance along with the magnitude of measurements. However, they are purely descriptive methods and cannot provide inference regarding agreement. A formal tool for assessing local agreement pattern would be desirable, but has been lacking in literature.

The objective of this work is to study the local agreement pattern between two continuous measurements subject to censoring (i.e. survival outcomes). Survival outcomes are frequently observed in biomedical studies and it may be of interest to assess the agreement between survival times measured on the same subjects using different methods. For example, in depression studies, the time of onset of clinical depression is measured using both clinician-administered Hamilton depression rating scale (HAM-D) and a patient self-report dimensional instrument (Carroll-D). Evaluating agreement between the disease onset times based on the different instruments is useful in assessing whether the less time-consuming and easier to use patient self-report instrument could be adopted as a reasonable replacement for the clinician-administered instrument. It is clear that the data setting considered here is more challenging than the typical case of complete observations of continuous measurements because the outcomes can be censored. Furthermore, as the measurements are observed over time, it is of interest to study how local agreement pattern changes over time.

Most existing agreement measures with censored survival data have been focused on global measures (Liu et al., 2005; Guo and Manatunga, 2007, 2010; Guo et al., 2013) and cannot be used to characterize local agreement patterns. The descriptive tools for agreement patterns, such as scatter plot and Bland and Altman plot, are not applicable for survival data due to the presence of censoring. Recently, Zeng et al. (2015) proposed a time-dependent agreement method to assess temporal pattern of agreement between two time-to-event endpoints. Their approach is likelihood-based and thus requires parametric distribution assumptions.

In this manuscript, we propose a local agreement measure for survival data to describe the local agreement pattern by considering the bivariate hazard function with an adjustment made to the expected chance agreement alone. It is sensible to consider hazard functions in this setting, due to its interpretability and analytical simplifications related to censoring. An appealing feature of our proposed local agreement measure is that it has a nice connection with the local kappa coefficient developed for assessing local agreement between two discrete survival times (Guo and Manatunga, 2005). Specifically, the proposed agreement measure can be viewed as an extension of the unscaled local kappa coefficient for continuous survival times. Our approach taken here is conceptually different from the modeling approaches suggested by Tanner and Young (1985), Agresti (1988), and Zeng et al. (2015), since we do not fit a series of models specified under specific types of dependence structures for describing the agreement pattern.

Our proposed local agreement measure is a new addition to the literature on local dependence measure for correlated survival times. The existing methods include the local conditional hazards ratio, also known as the cross ratio (Clayton, 1978; Oakes, 1989; Hu et al., 2011; Nan et al., 2006), the tail dependence coefficient, which describes the amount of extremal dependence (Kolev, 2006), the conditional Spearman's rho (Kotz and Nadarajah, 2002) and the local association measure in Cheng (2009). Comparing to our proposed local agreement measure, the existing local dependence measures mainly reflect the strength of association between bivariate outcomes and do not intend to capture the differences in the marginal distributions of the variables. In comparison, the local agreement measure in the current work aims to capture the strength of agreement between the two outcomes. As pointed out in Lin (1989), a distinct feature of an agreement measure is that it not only reflects the strength of association between two correlated outcomes but also reflects their marginal heterogeneity. To help illustrate this point, we show in the paper that the proposed local agreement measure is a function of the cross ratio, which reflects the strength of association and another term which could potentially captures the marginal heterogeneity via the pattern of the local single failure hazards. We also demonstrate their difference empirically by comparing the pattern of the local agreement measure and the cross ratio for bivariate survival data with different marginal distributions. Results show that the proposed agreement measure could be more effective than the cross ratio to capture the marginal heterogeneity.

For estimation of the proposed agreement measure, we develop a non-parametric approach which does not impose any assumptions on marginal survival distributions or the dependence structure between the two survival times. Hence, the local agreement pattern reflected by our proposed measure is not restricted by specific dependence structures. We use kernel methods (Fermanian, 1997) to estimate the hazard functions. Since survival outcomes are inherently bounded in the positive region, we propose to improve the estimation accuracy by adopting appropriate boundary kernel correction. The proposed estimators for the agreement measure have desirable asymptotic properties such as strong consistency and asymptotic normality. We perform simulation studies to evaluate the performance of the estimation method. We use a prostate cancer study to demonstrate the practical utility of our method.

2. Methods

2.1 Proposed local agreement pattern measure $\phi(t_1, t_2)$

In this section, we first present the proposed local agreement pattern measure and discuss its properties. To set notations, let T_1 and T_2 denote a pair of continuous survival times of the same individual based on different raters or methods. The joint survival function of the correlated survival times (T_1, T_2) is denoted as $S(t_1, t_2) = Pr(T_1 > t_1, T_2 > t_2)$. For any time point (t_1, t_2) within the support of the bivariate survival function, denote $(t_1, t_1 + t_1)$ and $(t_2, t_2 + t_2)$ as very small time intervals immediately following t_1 and t_2 . Given that $T_1 > t_1$ and $T_2 > t_2$, let's consider all possible conditional survival status of T_1 and T_2 within the rectangular region $(t_1 + t_1, t_2 + t_2)$ which include: $T_1 > t_1 + t_1, T_2 > t_2 + t_2$; $T_1 \in [t_1, t_1 + t_1), T_2 > t_2 + t_2$; $T_1 > t_1 + t_1, T_2 \in [t_2, t_2 + t_2)$; and $T_1 \in [t_1, t_1 + t_1), T_2 \in [t_2, t_2 + t_2)$. The corresponding conditional probabilities are defined as follows:

$$\begin{aligned} P_{00}(t_1, t_2) &= Pr(T_1 > t_1 + \Delta t_1, T_2 > t_2 + \Delta t_2 | T_1 > t_1, T_2 > t_2) = \frac{S(t_1 + \Delta t_1, t_2 + \Delta t_2)}{S(t_1, t_2)}, \\ P_{10}(t_1, t_2) &= Pr(T_1 \in [t_1, t_1 + \Delta t_1), T_2 > t_2 + \Delta t_2 | T_1 > t_1, T_2 > t_2) = \frac{S(t_1, t_2 + \Delta t_2) - S(t_1 + \Delta t_1, t_2 + \Delta t_2)}{S(t_1, t_2)}, \\ P_{01}(t_1, t_2) &= Pr(T_1 > t_1 + \Delta t_1, T_2 \in [t_2, t_2 + \Delta t_2) | T_1 > t_1, T_2 > t_2) = \frac{S(t_1 + \Delta t_1, t_2) - S(t_1 + \Delta t_1, t_2 + \Delta t_2)}{S(t_1, t_2)}, \\ \text{and } P_{11}(t_1, t_2) &= Pr(T_1 \in [t_1, t_1 + \Delta t_1), T_2 \in [t_2, t_2 + \Delta t_2) | T_1 > t_1, T_2 > t_2) \\ &= \frac{S(t_1, t_2) - S(t_1 + \Delta t_1, t_2) - S(t_1, t_2 + \Delta t_2) + S(t_1 + \Delta t_1, t_2 + \Delta t_2)}{S(t_1, t_2)}. \end{aligned}$$

We define the following measure to capture the chance-corrected local agreement between the two survival times within the rectangular region $[t_1, t_1 + t_1) \times [t_2, t_2 + t_2)$ as,

$$k(t_1 + \Delta t_1, t_2 + \Delta t_2) = P_{00}(t_1, t_2) + P_{11}(t_1, t_2) - P_{1+}(t_1, t_2)P_{+1}(t_1, t_2) - P_{0+}(t_1, t_2)P_{+0}(t_1, t_2),$$

(1)

where $P_{0+}(t_1, t_2) = P_{00}(t_1, t_2) + P_{01}(t_1, t_2)$, $P_{+0}(t_1, t_2) = P_{00}(t_1, t_2) + P_{10}(t_1, t_2)$, $P_{1+}(t_1, t_2) = P_{11}(t_1, t_2) + P_{10}(t_1, t_2)$, $P_{+1}(t_1, t_2) = P_{11}(t_1, t_2) + P_{01}(t_1, t_2)$ are the marginal probabilities in the 2×2 table. From (1), we note that $k(t_1 + t_1, t_2 + t_2)$ is closely related to a local Kappa coefficient proposed by Guo and Manatunga (2005) for measuring local agreement between discrete survival times. Specifically, the observed local agreement probability for the two survival outcomes in the 2×2 table is $P_{00}(t_1, t_2) + P_{11}(t_1, t_2)$. The expected agreement probability when the two survival outcomes are locally independent is $P_{1+}(t_1, t_2)P_{+1}(t_1, t_2) + P_{0+}(t_1, t_2)P_{+0}(t_1, t_2)$. Therefore, $k(t_1 + t_1, t_2 + t_2)$ is the unscaled version of the local Kappa coefficient (Guo and Manatunga, 2005) defined based on the 2×2 table representing the local region $[t_1, t_1 + t_1) \times [t_2, t_2 + t_2)$.

Motivated by (1), we propose a local agreement measure $\phi(t_1, t_2)$ by taking the limit of the regional chance-corrected agreement function $k(t_1 + t_1, t_2 + t_2)$ as the region area goes to 0. Specifically, $\phi(t_1, t_2)$ is defined as follows,

$$\varphi(t_1, t_2) = \frac{1}{2} \lim_{\Delta t_1, \Delta t_2 \rightarrow 0} \frac{k(t_1 + \Delta t_1, t_2 + \Delta t_2)}{\Delta t_1 \Delta t_2} = \lambda_{11}(t_1, t_2) - \lambda_{10}(t_1, t_2) \lambda_{01}(t_1, t_2), \quad (2)$$

where λ_{11} , λ_{10} and λ_{01} are bivariate hazard functions corresponding to local “double” and “single” failures (Dabrowska, 1988), which are defined as,

$$\begin{aligned} \lambda_{11}(t_1, t_2) &= \lim_{\Delta t_1, \Delta t_2 \rightarrow 0} \frac{\Pr(T_1 \in [t_1, t_1 + \Delta t_1), T_2 \in [t_2, t_2 + \Delta t_2) | T_1 \geq t_1, T_2 \geq t_2)}{\Delta t_1 \Delta t_2}, \\ \lambda_{10}(t_1, t_2) &= \lim_{\Delta t_1 \rightarrow 0} \frac{\Pr(T_1 \in [t_1, t_1 + \Delta t_1) | T_1 \geq t_1, T_2 \geq t_2)}{\Delta t_1}, \text{ and} \\ \lambda_{01}(t_1, t_2) &= \lim_{\Delta t_2 \rightarrow 0} \frac{\Pr(T_2 \in [t_2, t_2 + \Delta t_2) | T_1 \geq t_1, T_2 \geq t_2)}{\Delta t_2}. \end{aligned}$$

According to Equations (1) and (2), when there is no local agreement beyond expected by chance, our agreement measure $\varphi(t_1, t_2)$ equals 0. When the two survival outcomes have stronger agreement than expected by chance at (t_1, t_2) , $\varphi(t_1, t_2)$ has a positive value, with larger value indicating stronger local agreement. When the local agreement is weaker than the expected by chance, $\varphi(t_1, t_2)$ becomes negative. Since $\varphi(t_1, t_2)$ is defined based on the bivariate hazard functions which are not bounded above, $\varphi(t_1, t_2)$ is not restricted within a fixed range, which means it belongs to the category of unscaled agreement measures. Therefore, it is not advisable to interpret the magnitude of the proposed measure at a given time point against a fixed lower or upper bound. Instead, one should examine the pattern of proposed local agreement measures across the two dimensional time plane to investigate how the local agreement pattern changes across time for the correlated survival outcomes. Furthermore, given that the local agreement measure depends on marginal distributions, one should not directly compare the magnitude of the measure between two datasets with different marginals.

2.2 Properties of $\varphi(t_1, t_2)$ as a local agreement pattern measure

The proposed local agreement pattern measure $\varphi(t_1, t_2)$ has a connection with the cross ratio (Clayton, 1978; Oakes, 1989), a commonly used local dependence measure for continuous bivariate survival times. For any (t_1, t_2) within the support of the bivariate survival function, the cross ratio $\theta(t_1, t_2)$ is defined as,

$$\theta(t_1, t_2) = \frac{S(t_1, t_2) \frac{d^2}{dt_1 dt_2} S(t_1, t_2)}{\frac{d}{dt_1} S(t_1, t_2) \frac{d}{dt_2} S(t_1, t_2)}, \quad (3)$$

where $\theta(t_1, t_2)$ equals 1 when there is local independence between the two survival outcomes and is great than 1 when there is positive local dependence, with a higher value indicating stronger positive dependence. From the definitions, we can show that

$$\varphi(t_1, t_2) = \lambda_{10}(t_1, t_2) \lambda_{01}(t_1, t_2) (\theta(t_1, t_2) - 1). \quad (4)$$

Equation (4) suggests that the local agreement pattern measure $\varphi(t_1, t_2)$ changes in the same direction as the local dependence measure $\theta(t_1, t_2)$. That is, as the dependence between T_1 and T_2 increases, the local agreement increases. Also, $\varphi(t_1, t_2)$ becomes zero when there is local independence between the two survival times. Another insight from (4) is that as an agreement measure, φ not only depends on the association, reflected by θ , but also evaluates the homogeneity in the “single failure” hazards (Dabrowska, 1988), λ_{10} and λ_{01} . Specifically, keeping the sum of the two local “single failure” hazards constant, the product $\lambda_{10}\lambda_{01}$ increases as the absolute difference between λ_{10} and λ_{01} decreases. In other words, $\varphi(t_1, t_2)$ increases as the local single failure hazards become more homogenous. This result shows that the local agreement measure evaluates both the association as well as local homogeneity of the two survival outcomes, which is an essential feature for agreement measures (Lin, 1989; Lin et al., 2002). Furthermore, the product $\lambda_{10}(t_1, t_2)\lambda_{01}(t_1, t_2)$, which reflects the homogeneity in the local single failure hazards, is often related to the marginal homogeneity, in the sense that this product tends to have a more symmetric pattern around 45 degree line with more homogeneous marginals. Therefore, the pattern of the proposed local agreement measure could potentially reflect the degree of marginal homogeneity of bivariate survival times, which we will demonstrate later.

To illustrate graphically, how the pattern of this local agreement measure captures the strength of local dependence of bivariate survival times, we generate three random samples of bivariate survival times with the sample size of 500 from the Clayton model with the same marginal survival functions but with different cross ratios, which reflect different degrees of dependence. Next, we introduce 50% censoring to the bivariate survival times. In Fig 1a(a) and 1a(b), we plot the bivariate survival times with complete data and censored data, respectively. In Fig 1a(a), when the two survival times have homogeneous marginal distributions, they are mostly distributed along the 45 degree line with more concentration along this line when there is higher strength of dependence. However, when the survival times are subject to censoring (Fig 1a(b)), this pattern is less obvious and it becomes harder to infer the strength of dependence between the two survival outcomes based on the scatter plot of the observed survival times. In Fig 1a(c), we present the heatmaps representing the local agreement measure $\varphi(t_1, t_2)$ on the two dimensional time plane. The heatmaps of $\varphi(t_1, t_2)$ clearly demonstrate different patterns of the local agreement measure for survival times with different strengths of dependence. Specially, when the two survival times have the same marginal distributions and the highest dependence, the $\varphi(t_1, t_2)$ has the highest values along the 45 degree line, i.e. $t_1 = t_2$, and decreases fast as moving away from the 45 degree line, i.e. as $|t_1 - t_2|$ increases. When the two survival times are less dependent, $\varphi(t_1, t_2)$ is less elevated on the 45 degree line and decreases much slower as (t_1, t_2) moving away from the 45 degree line. As a comparison, we also present heatmaps of the cross ratio in Fig 1a(d), which remains constant across time in each scenario.

To demonstrate how the pattern of the $\varphi(t_1, t_2)$ can potentially capture the marginal heterogeneity between two survival times, we generate three samples of bivariate survival

times with the sample size of 500 from the Clayton model with the same cross ratio but with different marginal survival functions. Similarly, we introduce 50% censoring to the bivariate survival times. In Fig 1b(a), with heterogeneous marginal distributions, the bivariate survival times are no longer systemic around the 45 degree line and shifted further away from it with the increase of the heterogeneity. This pattern, however, is not obvious from the scatter plots with censored observations Fig 1b(b). Fig 1b(c) shows that the pattern of $\varphi(t_1, t_2)$ clearly captures this type of disagreement with the the pattern of the local agreement measure becomes asymmetric around the 45 degree line and the highest local agreement area moving away from the line as the marginal survival functions become different. In comparison, the cross ratio in Fig 1b(d) does not reflect the difference in the marginal distributions. The reason that the proposed local agreement measure can better capture the difference in marginals is because $\varphi(t_1, t_2)$ is a function of the cross ratio multiplied by the product of $\lambda_{10}(t_1, t_2)\lambda_{01}(t_1, t_2)$ (see Equation (4)). This product is often related to marginal homogeneity. For the three samples generated from Clayton models, the product $\lambda_{10}(t_1, t_2)\lambda_{01}(t_1, t_2)$ becomes more asymmetric around the 45 degree line when the marginals become more heterogeneous. Consequently, $\varphi(t_1, t_2)$ shows more asymmetry around the 45 degree line as compared with the cross ratio.

To further demonstrate the added value of the proposed measure over the cross-ratio, we also generate three samples of bivariate survival times from the Frank models with the same strength of association but with different degree of marginal heterogeneity. From Web Figure 1, we note the bivariate survival times are shifted away from the 45 degree line due to the heterogeneous marginals. From Web Figure 1(c), the pattern of the proposed $\varphi(t_1, t_2)$ clearly captures the disagreement between the survival times with the local $\varphi(t_1, t_2)$ become less symmetric around the 45 degree line as the heterogeneity between the marginals increases. In comparison, the disagreement is not so obviously reflected by cross ratio in Web Figure 1(d). We have had similar findings with other bivariate survival time models such as the Gumbel model (see Web Figure 2).

Another appealing feature of the proposed local agreement measure is that it can fully capture the dependence structure between the correlated survival times. As shown in the Proposition 1, the survival function $S(t_1, t_2)$ is jointly determined by the marginal survival functions $S_1(t_1)$, $S_2(t_2)$ and an integrable local agreement pattern measure function $\varphi(t_1, t_2)$.

Proposition 1—Let $S_1(t) = S(t, 0)$, $S_2(t) = S(0, t)$, and $\tau = (0, \tau_1) \times (0, \tau_2)$, where $\tau_1 = \sup\{t: S_1(t) > 0\}$ and $\tau_2 = \sup\{t: S_2(t) > 0\}$, we have

$$S(t_1, t_2) = S_1(t_1)S_2(t_2)\exp \left\{ \int_0^{t_1} \int_0^{t_2} \varphi(u_1, u_2) du_1 du_2 \right\}.$$

The detailed proof of Proposition 1 is provided in Web Appendix A.

3. Estimation and Inference

3.1 Estimation and inference for $\varphi(t_1, t_2)$

We develop a nonparametric estimation method for the proposed local agreement pattern measure in the presence of censoring via kernel smoothing techniques. Let (T_{i1}, T_{i2}) ($i = 1, \dots, n$) be independent and identically distributed pairs of survival times observed from n independent subjects, which are subject to bivariate censoring by a pair of independent random variables $C_i = (C_{i1}, C_{i2})$. The observed data consist of random vectors $(\tilde{T}_{i1}, \tilde{T}_{i2}, \delta_{i1}, \delta_{i2})$ ($i = 1, \dots, n$), where $\tilde{T}_{ij} = \min(T_{ij}, C_{ij})$ and $\delta_{ij} = I(T_{ij} \leq C_{ij})$ for $j = 1, 2$. We propose to estimate $\varphi(t_1, t_2)$ by plugging in kernel estimators of the hazard functions. That is, the proposed estimator is given by

$$\hat{\varphi}(t_1, t_2) = \hat{\lambda}_{11}(t_1, t_2) - \hat{\lambda}_{10}(t_1, t_2) - \hat{\lambda}_{01}(t_1, t_2),$$

where $\hat{\lambda}_{10}(t_1, t_2) = \int_0^{t_1} \frac{1}{h} K_1\left(\frac{v_1 - t_1}{h}\right) d\hat{\Lambda}_{10}(dv_1, t_2)$, $\hat{\lambda}_{01}(t_1, t_2) = \int_0^{t_2} \frac{1}{h} K_1\left(\frac{v_2 - t_2}{h}\right) d\hat{\Lambda}_{01}(t_1, dv_2)$, and $\hat{\lambda}_{11}(t_1, t_2) = \int_0^{t_1} \int_0^{t_2} \frac{1}{h^2} K_2\left(\frac{v_1 - t_1}{h}, \frac{v_2 - t_2}{h}\right) d\hat{\Lambda}_{11}(dv_1, dv_2)$. Here, $\hat{\Lambda} = (\hat{\Lambda}_{10}, \hat{\Lambda}_{01}, \hat{\Lambda}_{11})^T$ are the empirical estimators, for example the Nelson-Aalen estimators, of the corresponding cumulative hazard functions $\Lambda = (\Lambda_{10}, \Lambda_{01}, \Lambda_{11})^T$ (Dabrowska, 1988, 1989; Fermanian, 1997). Using notations similar to Dabrowska (1989), the cumulative hazard functions are defined as

$\Lambda_{10}(t_1, t_2) = -\int_0^{t_1} \frac{H_{10}(dv_1, t_2)}{H_{00}(v_1^-, t_2)}$, $\Lambda_{01}(t_1, t_2) = -\int_0^{t_2} \frac{H_{01}(t_1, dv_2)}{H_{00}(t_1, v_2^-)}$, $\Lambda_{11}(t_1, t_2) = \int_0^{t_1} \int_0^{t_2} \frac{H_{11}(dv_1, dv_2)}{H_{00}(v_1^-, v_2^-)}$, and $H_{00}(v_1, v_2) = \Pr(\tilde{T}_1 > v_1, \tilde{T}_2 > v_2)$, where $H_{10}(v_1, v_2) = \Pr(\tilde{T}_1 > v_1, \tilde{T}_2 > v_2, \delta_1 = 1)$, $H_{01}(v_1, v_2) = \Pr(\tilde{T}_1 > v_1, \tilde{T}_2 > v_2, \delta_2 = 1)$, and $H_{11}(v_1, v_2) = \Pr(\tilde{T}_1 > v_1, \tilde{T}_2 > v_2, \delta_1 = 1, \delta_2 = 1)$. K_1 is a one-dimensional kernel function which has support on $[-1, 1]$ with integral one and has bounded variation and bounded first derivative, K_2 is a two-dimensional kernel function with similar properties, and h is the smoothing parameter, which also known as the bandwidth. Here, we assume common bandwidth for the 3 kernels to facilitate the demonstration of the derivations of the theoretical results for the proposed estimator. When applying of the proposed method, one can adopt different bandwidths, such as adopting different bandwidth for estimating the single failure hazards and the double failure hazard. In Section 3.2, we will present more details on the choice of kernel functions and bandwidths. One can develop other estimators for the local hazards such as adopting alternative cumulative hazard function estimators (Peterson, 1977; Bantis et al., 2012).

Next, we establish the asymptotic properties of the proposed estimator. To facilitate the following derivation, we first define an alternative formulation for the proposed agreement pattern measure and its estimator. Denote $\lambda(\cdot) = (\lambda_{10}(\cdot), \lambda_{01}(\cdot), \lambda_{11}(\cdot))^T$. Let λ_θ be the collection of bivariate hazard functions λ on R_2 . Define the functional $g: \lambda_\theta \mapsto R$ as follows:

$$g(\lambda) = e_3^T \lambda - e_1^T \lambda e_2^T \lambda, \quad (5)$$

where $\mathbf{e}_1 = (1, 0, 0)^T$, $\mathbf{e}_2 = (0, 1, 0)^T$, $\mathbf{e}_3 = (0, 0, 1)^T$. It is straightforward to show the proposed local agreement pattern measure $\varphi(\mathbf{t})$, $\mathbf{t} = (t_1, t_2)$ can be expressed as a function of the bivariate hazard functions $\lambda(\cdot)$ via $g(\cdot)$, i.e., $\varphi(\mathbf{t}) = g(\lambda(\mathbf{t}))$. Our proposed estimator in (5) can be equivalently expressed as

$$\hat{\varphi}(\mathbf{t}) = g(\hat{\lambda}(\mathbf{t})), \quad (6)$$

where $\hat{\lambda} = (\hat{\lambda}_{10}(\cdot), \hat{\lambda}_{01}(\cdot), \hat{\lambda}_{11}(\cdot))^T$ are the kernel estimators of the hazard functions. In the following, we provide the asymptotic properties of $\hat{\varphi}$.

Theorem 1—Consider a region $\tau = (0, \tau_1) \times (0, \tau_2)$ where (τ_1, τ_2) is in the support of the observed event times. Assuming the bivariate hazard function estimator $\hat{\lambda}$ is uniformly strongly consistent within τ . For any $\mathbf{t} = (t_1, t_2) \in \tau$, the proposed estimator $\hat{\varphi}(\mathbf{t})$ has the following asymptotic properties as $n \rightarrow \infty$,

- i. The estimator $\hat{\varphi}(\mathbf{t})$ is strongly consistent, i.e., $|\hat{\varphi}(\mathbf{t}) - \varphi(\mathbf{t})| \rightarrow 0$ with probability 1.
- ii. The proposed estimator $\hat{\varphi}(\mathbf{t})$ has the following weak convergence result,

$$r_n \{\hat{\varphi}(\mathbf{t}) - \varphi(\mathbf{t})\} \xrightarrow{d} g'_\lambda(\mathbf{W}(\mathbf{t})),$$

where $r_n = (nh^2)^{1/2}$, $\mathbf{W}(\cdot)$ is a multivariate zero-mean Gaussian process with the covariance function defined in equation (2.10) of Fermanian (1997), and g'_λ is the Hadamard derivative of g at λ defined as

$$g'_\lambda(\mathbf{W}) = \mathbf{e}_3^T \mathbf{W} - \mathbf{e}_1^T \mathbf{W} \cdot \mathbf{e}_2^T \lambda - \mathbf{e}_2^T \mathbf{W} \cdot \mathbf{e}_1^T \lambda.$$

Here, $g'_\lambda(\mathbf{W}(\mathbf{t}))$ follows a zero-mean normal distribution.

- iii. By randomly sampling with replacement from the observed data $(\tilde{T}_{i1}, \tilde{T}_{i2}, \delta_{i1}, \delta_{i2})$ ($i = 1, \dots, n$), a bootstrap estimator $\hat{\varphi}^\#(\mathbf{t})$ can be obtained based on the bootstrap samples. Then $r_n \{\hat{\varphi}^\#(\mathbf{t}) - \hat{\varphi}(\mathbf{t})\}$, given the observed data, weakly converges to the same limiting distribution as $r_n \{\hat{\varphi}(\mathbf{t}) - \varphi(\mathbf{t})\}$ in probability.

We prove Theorem 1 based on the uniform consistency of kernel hazard function estimators $\hat{\lambda}$, the Hadamard differentiability of functional g and the functional delta method. The detailed proof of this theorem is provided in Web Appendix B. The assumption that the bivariate hazard function estimator $\hat{\lambda}$ is uniformly strongly consistent within τ has been shown to be valid under some regularity conditions of the bandwidths (Fermanian, 1997). Due to the complexity of the covariance function of $\hat{\lambda}$, an explicit expression for the asymptotic variance of $\hat{\varphi}(t_1, t_2)$ is analytically complicated. Theorem 1(iii) suggests an alternative approach is to use a bootstrap procedure to estimate the variance.

3.2 Choice of the kernel functions and bandwidths for estimating the hazard functions

In this section, we provide some discussions on the specifications of the kernel functions and bandwidths. We use the Epanechnikov kernels for estimating the hazard functions because they have been shown to be the most efficient kernel in minimizing the mean integrated squared error (Wang and Jones, 1995) and they also provide computational advantages over other kernel functions. For estimating $\lambda_{10}(t_1, t_2)$ and $\lambda_{01}(t_1, t_2)$, we adopt the univariate Epanechnikov kernel defined as $K_1(u) = 3/4(1 - u^2)I[u_2 - 1]$. For estimating $\lambda_{11}(t_1, t_2)$, we apply the bivariate product Epanechnikov kernel $K_2(u_1, u_2) = K_1(u_1)K_1(u_2)$.

One important consideration in kernel hazard rate estimation is that the hazard functions have bounded support. In univariate hazard rate estimation, boundary bias occurs when the support of the kernel function at a time point within the interval $[0, h]$ exceeds the available range of the observed data, and thus leads to increased bias. For bivariate survival data, boundary effects are observed when either T_1 or T_2 is close to zero. To reduce the boundary effects in hazard rate estimation, Müller and Wang (1994) proposed a class of boundary kernel functions which have shown numerical benefits in terms of smaller asymptotic mean squared error when estimating near the boundaries than other boundary kernels. Specifically, for estimating univariate hazard function in the region $B_L = \{t: 0 \leq t < h\}$, they proposed the

boundary kernel as $K_{1,t}(u) = \frac{12}{(1-q)^4}(u+1) [u(1-2q) + (3q^2 - 2q + 1)/2]$, where $q = t/h$ and $u \in [-1, q]$. In this paper, we extend Müller and Wang (1994)'s univariate boundary kernel to the bivariate case. For boundary regions $B_{L,I} = [t_1, t_2: 0 \leq t_1 < h, t_2 \leq h]$, $B_{I,L} = [t_1, t_2: t_1 > h, 0 \leq t_2 \leq h]$, and $B_{L,L} = [t_1, t_2: 0 \leq t_1 < h, t_2 \leq h]$, the proposed bivariate boundary kernel is formulated as $K_{2,t_1,1}(u_1, u_2) = K_{1,t_1}(u_1)K_1(u_2)$, $K_{2,1,t_2}(u_1, u_2) = K_1(u_1)K_{1,t_2}(u_2)$, and $K_{2,t_1,t_2}(u_1, u_2) = K_{1,t_1}(u_1)K_{1,t_2}(u_2)$ for $(t_1, t_2) \in B_{L,L}$. Here, $q_1 = t_1/h$, $q_2 = t_2/h$, $u_1 \in [-1, \min(1, q_1)]$ and $u_2 \in [-1, \min(1, q_2)]$.

Another consideration in our kernel estimation is the selection of bandwidths for estimating the bivariate hazard functions. Various bandwidth selection methods have been proposed for multivariate kernel density estimation (Wang and Jones, 1995; Duong and Hazelton, 2003), but very few have been proposed for multivariate hazard rate estimation. Fermanian (1997) proposed an asymptotically optimal plug-in bandwidth for estimating multivariate hazard functions. However, this bandwidth method typically requires a very large sample size and hence is not applicable in many studies with small to moderate sample sizes (less than several thousands). According to Fermanian (1997), a practical choice of bandwidth is to use Silverman's rule or Scott's rule (Silverman, 1986; Scott, 1992). In the simulation studies and data application, we choose the bandwidths for kernel estimation based on Scott's rule given by $h = n^{-1/(4+d)} \hat{\sigma}$, where d is the number of dimensions of the hazard function and $\hat{\sigma}$ is the sample standard deviation of the observed survival times.

4. Simulation Studies

We conducted simulation studies to assess the performance of the proposed estimation and inference procedure of the new local agreement pattern measure. In each simulation, we generated bivariate survival times (T_1, T_2) from the Clayton model (Clayton, 1978) with a sample size of 500. We considered two sets of simulation studies with different setups for

marginal distributions. In the first setup, we assumed T_1 and T_2 had identical marginal distributions which were standard exponential. In the second setup, the two survival times had different marginal distributions where T_1 and T_2 had exponential distributions with the means equal to 1 and 1.5, respectively. We specified a cross ratio of 3 in the Clayton model which indicated moderate dependence between bivariate survival times. The survival times generated from the Clayton model were subject to independent right censoring by two independent and exponentially distributed censoring variables. We considered three censoring rates of 17%, 33% and 50%, representing light, medium and heavy censoring, respectively.

Table 1 summarizes the results based on 500 simulation runs under various simulation scenarios. We selected nine time points (t_1, t_2) on the two dimensional plane where the values of $t_j (j = 1, 2)$ were chosen as the 10th, 30th and 50th percentiles of the standard exponential distribution. For each pair of time points, we presented the true value for the local agreement measure $\phi(t_1, t_2)$, empirical mean and empirical standard error for $\hat{\phi}(t_1, t_2)$, along with the average standard error estimate. We also presented coverage probability for the 95% confidence intervals based on the 200 bootstrap samples. The proposed nonparametric estimator for local agreement pattern measure demonstrated reasonable accuracy with the bias being less than 10% at most time points. The empirical standard error of the proposed estimator decreased as the censoring proportion decreased. The bootstrap standard error was close to the Monte Carlo standard error and the coverage probability of the estimated 95% confidence intervals was close to the nominal level in most cases.

To further illustrate the performance of the proposed estimator, we plotted the estimated local agreement pattern measures along with the true values over the two-dimensional time plane for identical (Fig 2a) and different (Fig 2b) marginal distributions. Specifically, the top panels of Figure 2a and 2b present the surfaces of the true local agreement pattern measure on the two-dimensional time plane. In the bottom panels, we first fixed T_1 at various values and then plotted the true and estimated profiles of ϕ across T_2 ranging from 0.05 to 1.45 (which was approximately 5% to 75% quantile of the standard exponential distribution). We also plotted the 95% pointwise Monte Carlo confidence intervals based on the empirical variance of $\hat{\phi}$. Then we fixed T_2 and plotted the profiles of ϕ across T_1 . Results from Figure 2a and 2b suggest that the proposed estimator provided fairly accurate estimation of the local agreement pattern measures across various time points in the two-dimensional time plane. The bias was in general small with the largest bias observed in the boundary region when either T_1 or T_2 was very close to 0. The Monte Carlo confidence bands for ϕ provided a nice coverage of the profile of the local agreement measure. The confidence bands tended to be wider near the boundary regions as expected. The simulation results show that the width of the confidence bands of the proposed local agreement measure estimator did not increase significantly near the boundary regions indicating reasonable performance of the estimator near the boundary regions.

5. Prostate Cancer Data

We illustrate the application of the proposed method using the data from a prostate cancer study. Various kinds of treatments are available for prostate cancer and it is of interest to

compare the efficacy of different treatments. One major difficulty is that physicians have been applying different definitions of disease-free status for specific treatments. Hence, the relapse-free survival rates for different treatments are based on their corresponding definitions and potential discrepancies between the definitions may cause misleading conclusions on treatment efficacy. For example, radical prostatectomy and irradiation are two commonly used treatments for prostate cancer. Different definitions have been proposed for the two treatments. For radical prostatectomy, post-treatment disease freedom is defined by reaching and maintaining an undetectable prostate specific antigen (PSA) nadir ranging between 0.2 and 0.5 ng/ml (Critz et al., 1996). For irradiation, disease freedom is represented by a non-rising PSA with the increasing PSA defined as three consecutive PSA increases measured 6 months apart, according to the American Society of Therapeutic Radiation Oncology (ASTRO) consensus criteria (1997). In order to accurately compare the relapse-free survival rates between the two treatments, it is important to first assess the agreement between the two disease-free definitions. In particular, we are interested in finding out how the agreement between the two definitions evolves along the time after the treatments.

In a clinical study, 1369 men received simultaneous radiotherapy for prostate cancer followed by an external beam radiation. The disease-free status was evaluated frequently after radiation treatment. The relapse-free time was defined as the time from the end of the irradiation till the prostate cancer relapse based on two different definitions. T_1 was the observed relapse-free time with disease recurrence defined as post-treatment PSA level exceeding the nadir of 0.2 ng/ml. T_2 was defined based on ASTRO definition and represented the midpoint between the time when the lowest PSA was achieved after irradiation and the time when the first of three consecutive rises in the PSA level occurred. The relapse-free times for a patient were subject to independence censoring due to the end of the follow-up on this patient. Among the 1369 patients, 159 subjects were diagnosed with prostate cancer recurrence according to both definitions and 64 had relapses based on only one of the definitions, indicating approximate 80% of censoring. Figure 3 presents the distribution of patients' observed relapse-free times measured by the two definitions. The plot shows that almost all of the observed cancer relapses happened within 8 years after the irradiation. Figure 3 also shows that most observed survival times were on the 45 degree line.

We applied the proposed measure ϕ to evaluate the local agreement pattern between T_1 and T_2 within 7.5 years after the irradiation. The estimated local agreement pattern measures are presented in Figure 4. On the top panel, we display the local agreement measure surface within the time space of $[0, 7.5] \times [0, 7.5]$. From the surface plot, we can see that within the early years after the treatment, the local agreement was highest along the 45 degree line and decreased dramatically when moving away from the 45 degree line. As time went by, the highest local agreement region moved slightly away from the 45 degree line towards $T_2 > T_1$ area. This suggest that if a patient remained disease-free for a while after irradiation, he was more likely to have a disease relapse diagnosed earlier by the nadir definition. On the bottom panel, we present the estimated local agreement pattern measures on the 45 degree line within 7.5 years and the corresponding 95% bootstrap pointwise confidence bands based on 1000 bootstrap samples. Note that we can only conduct pointwise inference for the local

agreement measure based on this bootstrap confidence band. Based on the figure, we can see that the local agreement pattern measure on the 45 degree line was highest around year 2 after the irradiation, suggesting the nadir and ASTRO definitions agreed best for cancer recurrences that happened around year 2 after the treatment. The local agreement decreased after year 2 indicating more disagreement between the two definitions for recurrences that happened beyond 2 years after the irradiation. This result was supported by our observations from Figure 3.

Guo and Manatunga (2007) had assessed the global strength of agreement for the prostate cancer data example using a nonparametric estimator of concordance correlation coefficient (CCC). The results based on the proposed local agreement measure in this paper provided new insights on how the strength of agreement between the two disease free survival definition changes over time after the treatment, i.e. the agreement between the two definitions decreased over time after patients survived beyond 2 years after the treatment.

6. Discussion

The proposed local agreement pattern measure is not bounded by a fixed range due to the nature of hazard functions being unbounded. As done with many other unbounded descriptive measures, such as cross ratio and local dependence function (Holland and Wang, 1987), the proposed measure can be interpreted with respect to the relative scale over the two-dimensional time space within the region of interest. It helps to address the nature of agreement when the relationship of two survival times is time dependent and can potentially be used for modeling such local relationship.

Compared to other dependence measure such as cross ratio, one distinct feature of the proposed an agreement measure is that it not only reflects the strength of association between the two correlated measurement but also capture the difference in their marginal distributions. This is because agreement measures are usually applied for assessing agreement on measurements taken on the same subjects using different methods or by different rater. Therefore, depending on the marginals is an intrinsic and necessary feature of an agreement measure (Lin, 1989; Lin et al., 2002). If the study goal is to investigate the association between two correlated outcomes instead of the difference/or agreement between them, an association measure such as cross ratio or Kendalls tau would be sufficient and there is no need to consider the marginal homogeneity in that case.

In the literature, there are some other related measures such as time-dependent concordance (Guo and Manatunga, 2007) and time-dependent accuracy measures (Cheng, 2015). These measures are closely related to the proposed local agreement measure and can be interpreted in a different way from the agreement.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the editor, the associate editor and the two referees for their valuable comments and suggestions. Research reported in this publication was supported by the National Institute Of Mental Health of the National Institutes of Health under Award Number R01 MH079448 and R01MH105561. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

- Agresti A. A model for agreement between ratings on an ordinal scale. *Biometrics*. 1988; 44:539–548.
- Agresti A. An agreement model with kappa as parameter. *Statistics & Probability Letters*. 1989; 7:271–273.
- Banerjee M, Capozzoli M, McSweeney L, Sinha D. Beyond kappa: A review of interrater agreement measures. *Canadian Journal of Statistics*. 1999; 27:3–23.
- Bantis LE, Tsimikas JV, Georgiou SD. Survival estimation through the cumulative hazard function with monotone natural cubic splines. *Lifetime Data Analysis*. 2012; 18(3):364–96. [PubMed: 22399231]
- Borg J, Mollgaard A, Riis BJ. Single x-ray absorptiometry: performance characteristics and comparison with single photon absorptiometry. *Osteoporosis International*. 1995; 5:377–381. [PubMed: 8800788]
- Cheng Y, Fine JP, Kosorok MR. Nonparametric Association Analysis of Exchangeable Clustered Competing Risks Data. *Biometrics*. 2009; 65:385–393. [PubMed: 18549422]
- Cheng Y, Li J. Time-dependent diagnostic accuracy analysis with censored outcome and censored predictor. *Journal of Statistical Planning and Inference*. 2015; 156:90–102.
- Clayton DG. A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika*. 1978; 65:141–151.
- Cohen J. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*. 1960; 20:37–46.
- Cohen J. Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin*. 1968; 70:213220.
- Critz FA, Levinson K, Williams WH, Holladay DA. Prostate-specific antigen Nadir: the optimum level after irradiation for prostate cancer. *Journal of Clinical Oncology*. 1996; 14:2893–2900. [PubMed: 8918485]
- Dabrowska DM. Kaplan-Meier Estimate on the Plane. *The Annals of Statistics*. 1988; 16:1475–1489.
- Dabrowska DM. Uniform Consistency of the Kernel Conditional Kaplan-Meier Estimate. *The Annals of Statistics*. 1989; 17(3):1157–1167.
- Darroch JN, McCloud PI. Category Distinguishability and Observer Agreement. *Australian Journal of Statistics*. 1986; 28:371–388.
- Duong T, Hazelton ML. Plug-in bandwidth matrices for bivariate kernel density estimation. *Journal of Nonparametric Statistics*. 2003; 15:17–30.
- Fermanian JD. Multivariate Hazard Rates under Random Censorship. *Journal of Multivariate Analysis*. 1997; 62:273–309.
- Guo Y, Manatunga AK. Modeling the Agreement of Discrete Bivariate Survival Times using Kappa Coefficient. *Lifetime Data Analysis*. 2005; 11:309–332. [PubMed: 16133882]
- Guo Y, Manatunga AK. Nonparametric estimation of the concordance correlation coefficient under univariate censoring. *Biometrics*. 2007; 63:164–72. [PubMed: 17447941]
- Guo Y, Manatunga AK. A note on assessing agreement for frailty models. *Statistics and Probability Letters*. 2010; 80:527–533. [PubMed: 20607108]
- Guo Y, Li R, Peng L, Manatunga AK. A new agreement measures based on survival processes. *Biometrics*. 2013; 69:874–882. [PubMed: 23844617]
- Holland PW, Wang YJ. Dependence function for continuous bivariate densities. *Communications in Statistics - Theory and Methods*. 1987; 16:863–876.

- Hu T, Nan B, Lin X, Robins JM. Time-dependent cross ratio estimation for bivariate failure times. *Biometrika*. 2011; 98:341–354. [PubMed: 22822258]
- Jones MC. The local dependence function. *Biometrika*. 1996; 83:899–904.
- Jung M, Dimtchev A, Velen A, Dritschilo A. Combining radiation therapy with interstitial radiation-inducible TNF-[alpha] expression for locoregional cancer treatment. *Cancer Gene Therapy*. 2011; 18(3):189–195. [PubMed: 21052099]
- Kolev N, Anjos U, Mendes B. Copulas: A review and recent developments. *Stochastic Models*. 2006; 22:617–660.
- Kotz K, Nadarajah D. Some local dependence functions for the elliptically symmetric distributions. *Sankhyā A*. 2002:65.
- Lin LI. A concordance correlation coefficient to evaluate reproducibility. *Biometrics*. 1989; 45:255–268. [PubMed: 2720055]
- Lin LI, Hedayat AS, Sinha B, Yang M. Statistical methods in assessing agreement: Models, issues and tools. *Journal of American Statistical Association*. 2002; 97(457):257–270.
- Liu X, Du Y, Teresi J, Hasin DS. Concordance correlation in the measurements of time to event. *Statistics In Medicine*. 2005; 24:1409–1420. [PubMed: 15568183]
- Müller HG, Wang JL. Hazard Rate Estimation under Random Censoring with Varying Kernels and Bandwidths. *Biometrics*. 1994; 50:61–76. [PubMed: 8086616]
- Nan B, Lin X, Lisabeth LD, Harlow S. Piecewise constant cross-ratio estimation for association of age at a marker event and age at menopause. *Journal of American Statistical Association*. 2006; 101:65–77.
- Oakes D. Bivariate survival models induced by frailties. *Journal of the American Statistical Association*. 1989; 84:48793.
- Peterson AV Jr. Expressing the Kaplan-Meier estimator as a function of empirical sub-survival functions. *Journal of the American Statistical Association*. 1977; 72:854–858.
- Prentice RL, Cai J. Covariance and survival function estimation using censored multivariate failure time data. *Biometrika*. 1992; 79:495–512.
- Schild RL, Fimmers R, Hansmann M. Fetal weight estimation by three-dimensional ultrasound. *Ultrasound in Obstetrics & Gynecology*. 2000; 16:445–452. [PubMed: 11169329]
- Scott, DW. *Multivariate Density Estimation: Theory, Practice, and Visualization*. John Wiley & Sons; New York, Chichester: 1992.
- Silverman, BW. *Density Estimation for Statistics and Data Analysis*. Chapman & Hall; London: 1986.
- Steven, PM. *EnvStats: An R Package for Environmental Statistics*. Springer; New York: 2013.
- Tanner MA, Young MA. Modeling ordinal scale disagreement. *Psychological Bulletin*. 1985; 98(2): 408–415. [PubMed: 3901069]
- van der Vaart, AW., Wellner, JA. *Weak Convergence and Empirical Processes*. Springer-Verlag; New York: 1996.
- Wand, MP., Jones, MC. *Kernel Smoothing*. Chapman & Hall; London: 1995.
- Zeng D, Cornea E, Dong J, Pan J, Ibrahim JG. Assessing temporal agreement between central and local progression-free survival times. *Statistics in Medicine*. 2015; 34:844–858. [PubMed: 25393731]

Figure 1a.

Figure 1b.

Figure 1.

Figure 1a. Clayton models with the same marginal exponential distributions for T1 and T2 but different dependence levels. The dependence between T1 and T2 decreases from left to right with Kendall's tau varying from 0.71, 0.62 to 0.5. From upper panel to lower panel, the figures represent the scatterplots with complete data, scatterplots with censored data, local agreement pattern measure heatmaps, and cross ratio heatmaps, respectively.

Figure 1b. Clayton models with the same strength of association and different degree of heterogeneity in the marginal distributions, where the heterogeneity increases from the left to the right columns. Specially, T1 has the exponential distribution of rate parameter of 1 in all three columns while T2 has the exponential distribution with the rate parameter taking values of 1, 0.6 and 0.3 for the left, middle and right columns, respectively. From upper panel to lower panel, the figures represent (a) the scatterplots with complete data generated from the Clayton models, (b) scatterplots with censored data, (c) the heatmaps of the proposed local agreement measure, (d) the heat maps of the local cross ratio.

Figure 2a.

Figure 2b.

Figure 2.

Figure 2a. Local agreement pattern measure for the Clayton model with homogeneous marginal distributions. The surface plot in the top panel is the true local agreement pattern measure on the 2-D time space. Line curves in the bottom panels correspond to the lines in the surface plot when fixing one of the two survival times. Dot curves are the estimated local agreement pattern measures and the corresponding empirical pointwise 95% confidence bands.

Figure 2b. Local agreement pattern measure for the Clayton model with heterogeneous marginal distributions. The surface plot in the top panel is the true local agreement pattern measure on the 2-D time space. Line curves in the bottom panels correspond to the lines in the surface plot when fixing one of the two survival times. Dot curves are the estimated local agreement pattern measures and the corresponding empirical pointwise 95% confidence bands.

Figure 3.

Prostate cancer relapse-free survival times after irradiation based on the two definitions of disease-free state. The figure on the upper panel presents the empirical joint distribution of T1 and T2. Red dots represent subjects with both survival times observed. The orange plus signs are corresponding to patients that did not experience disease relapse during the study according to both definitions. The blue arrow signs stand for patients that had been diagnosed with disease relapse with only one definition in the study. The figure on the lower panel presents the Kaplan Meier survival curves for T1 and T2.

Figure 4. Estimated local agreement pattern measure surface (top) and diagonal-line curve with 95% bootstrap pointwise confidence bands (bottom).

Table 1

Summary statistics of the proposed local agreement pattern measure estimator $\hat{\phi}(t_1, t_2)$ for the Clayton family with identical or different marginal distributions and Kendall's tau=0.5.

Marginal Distributions	T_1	Censoring	T_2		
			10%	30%	50%
Homogeneous	10%	Light	1.550 (0.280,0.290,92.2%)	1.046 (0.169,0.165,90.0%)	0.596 (0.164,0.147,91.0%)
		Medium	1.395 1.552 (0.281,0.296,91.4%)	0.967 1.050 (0.179,0.174,90.4%)	0.556 0.604 (0.178,0.162,91.2%)
		Heavy	1.549 (0.298,0.305,92.6%)	1.051 (0.187,0.191,92.2%)	0.605 (0.219,0.194,89.0%)
	30%	Light	1.050 (0.152,0.166,94.0%)	0.840 (0.096,0.097,92.4%)	0.622 (0.096,0.097,93.4%)
		Medium	0.967 1.046 (0.163,0.175,93.6%)	0.873 0.844 (0.106,0.104,92.4%)	0.646 0.620 (0.106,0.107,92.4%)
		Heavy	1.049 (0.174,0.190,95.4%)	0.842 (0.119,0.116,91.4%)	0.620 (0.123,0.127,92.6%)
	50%	Light	0.602 (0.144,0.148,93.8%)	0.619 (0.095,0.096,92.4%)	0.609 (0.100,0.106,92.4%)
		Medium	0.556 0.601 (0.154,0.162,95.2%)	0.646 0.621 (0.108,0.107,93.8%)	0.654 0.608 (0.116,0.119,92.4%)
		Heavy	0.603 (0.197,0.193,93.0%)	0.616 (0.128,0.128,93.0%)	0.609 (0.148,0.146,90.4%)
Heterogeneous	10%	Light	1.227 (0.269,0.254,90.2%)	0.885 (0.158,0.149,91.2%)	0.591 (0.138,0.138,94.8%)
		Medium	1.085 1.225 (0.268,0.258,91.4%)	0.833 0.883 (0.165,0.157,92.8%)	0.562 0.586 (0.155,0.153,94.2%)
		Heavy	1.224 (0.274,0.265,91.4%)	0.887 (0.181,0.172,92.0%)	0.601 (0.187,0.183,94.2%)
	30%	Light	0.800 (0.137,0.144,94.6%)	0.668 (0.083,0.084,94.2%)	0.562 (0.087,0.087,92.2%)
		Medium	0.728 0.798 (0.144,0.151,94.4%)	0.689 0.669 (0.090,0.090,92.8%)	0.579 0.559 (0.099,0.097,91.8%)
		Heavy	0.803 (0.162,0.165,94.0%)	0.670 (0.102,0.101,92.2%)	0.556 (0.116,0.114,92.0%)
	50%	Light	0.442 (0.124,0.125,94.4%)	0.473 (0.087,0.083,93.8%)	0.481 (0.090,0.091,92.6%)
		Medium	0.407 0.445 (0.138,0.138,94.0%)	0.465 0.469 (0.096,0.092,92.6%)	0.501 0.486 (0.100,0.103,93.4%)
		Heavy	0.437 (0.154,0.163,94.4%)	0.473 (0.110,0.110,93.4%)	0.477 (0.124,0.125,94.2%)