Quantifying the Average of the Time-varying Hazard Ratio via a Class of Transformations

Joseph G. Ibrahim

Department of Biostatistics, UNC

Joint work with Qingxia Chen, Donglin Zeng, Ming-Hui Chen, Zhiying Pan, Xiaodong Xue
ibrahim@bios.unc.edu

UPENN Clinical Trials

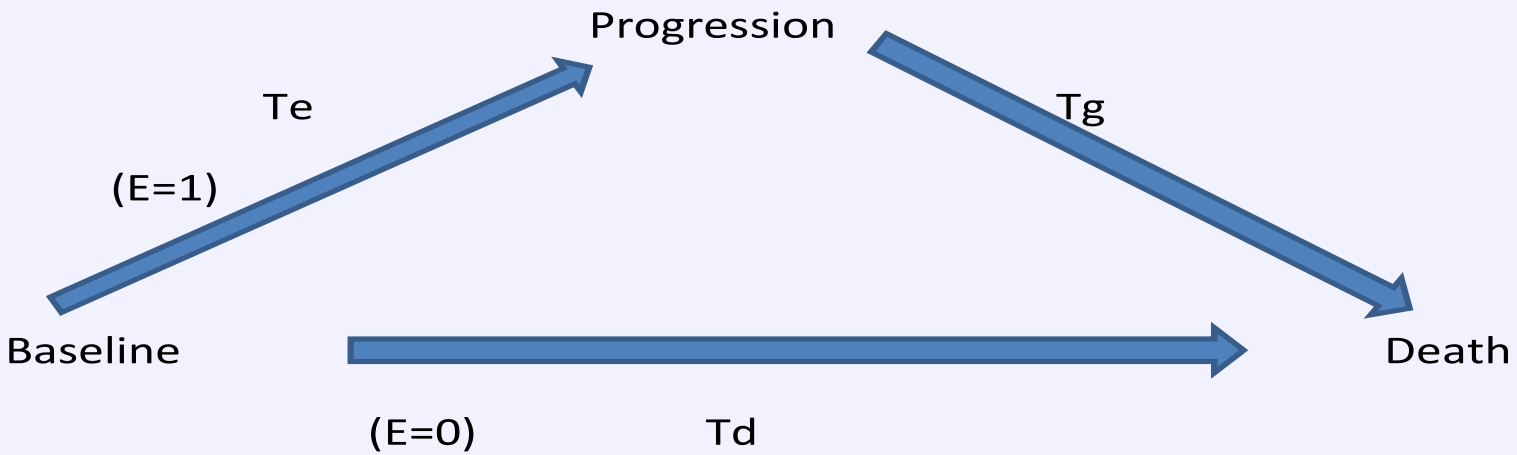
Outline

ICH Guidelines

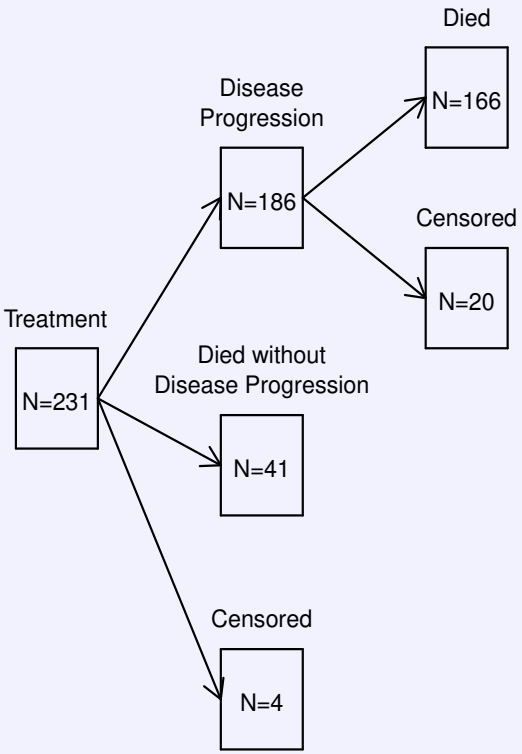
- The ICH guidelines on sensitivity and analysis in clinical trials address defining proper estimands and methodology in the presence of
  - treament switching
  - rescue medications
  - treatment discontinuation
  - terminal events

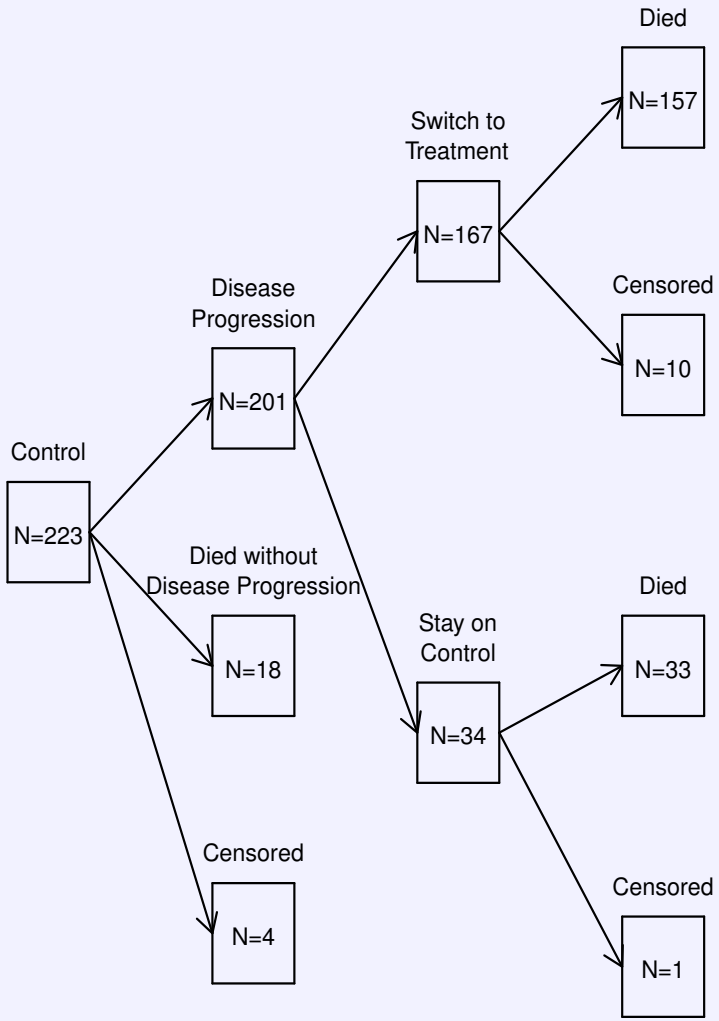
Motivation: Panitumumab Data (treatment switching)

- A phase III multi-center clinical trial with patients randomized to receive panitumumab plus best supportive care or best supportive care alone.
- Primary endpoint is overall survival
- During the trial, the patients receiving best supportive care alone were allowed to switch to the experimental treatment if they experience disease progression.
- In order to properly adjust for treatment switching, Zeng et al. (2012) propose a *Transition Model* (TM) with four submodels (similar to illness-death model):



Graphical Representation of the Panitumumab Data





Covariates for Panitumumab Data

- treatment
- Age
- gender
- bECOG: baseline electrocorticography performance status
- CenEastEU: central Europe, WesternEU: western Europe
- Prog Time: progression time, PR: partial response
- BTR: best tumor response, SD: stable disease
- LECOG: last electrocorticography performance status
- V: binary treatment switching covariate
- AE: binary adverse event covariate.
- Rectal: binary covariate of whether they had rectal or colon cancer

Table 1: Analysis of the Panitumumab Data

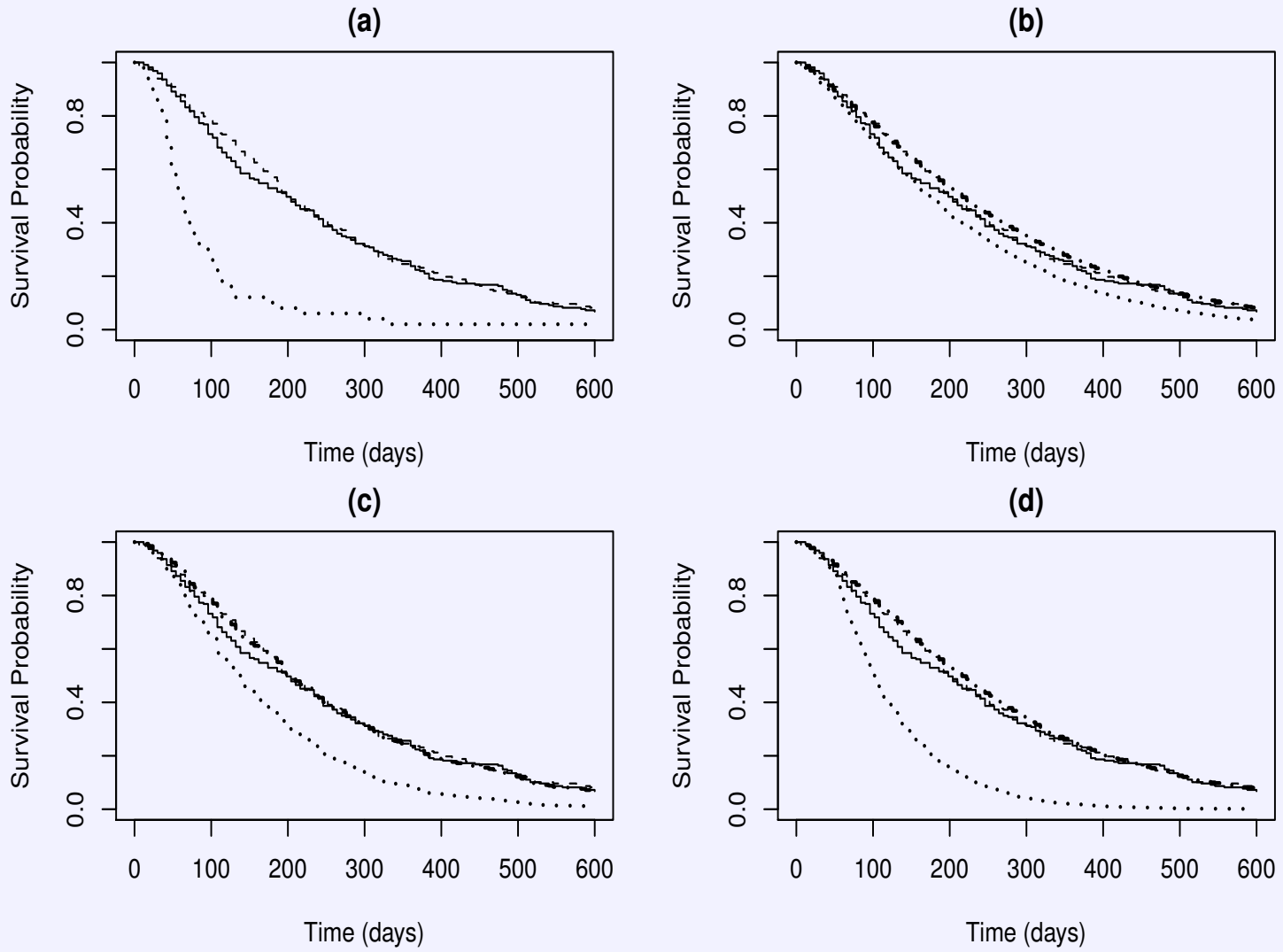
Parameter	EST	SD	P-value	Parameter	EST	SD	P-value
		T_D Model				T_E Model	
Treatment	-0.464	0.347	0.182	Treatment	-1.144	0.118	<.001
Age	0.023	0.015	0.124	Age	-0.015	0.005	0.004
bECOG01	-0.589	0.299	0.048	bECOG01	-0.805	0.174	<.001
Rectal	-0.028	0.320	0.929	Rectal	-0.018	0.110	0.871
Male	-0.288	0.305	0.345	Male	-0.054	0.109	0.622
CenEastEU	-0.188	0.627	0.764	CenEastEU	0.194	0.250	0.439
WesternEU	0.181	0.399	0.650	WesternEU	-0.068	0.160	0.672
		T_G Model				E Model	
Treatment	-0.784	0.214	<.001	Intercept	1.366	0.972	0.160
V*(1-Treatment)	-1.383	0.209	<.001	Treatment	-1.070	0.319	<.001
Prog Time	-0.003	0.001	0.039	Age	-0.008	0.014	0.546
PR Age	-0.004	0.005	0.450	bECOG01	1.905	0.334	<.001
BTR PR	-0.226	0.345	0.512	Rectal	0.314	0.331	0.342
BTR SD	-0.180	0.174	0.302	Male	-0.303	0.321	0.346
bECOG01	-0.268	0.196	0.173	CenEastEU	0.078	0.623	0.901
LECOG01	-1.035	0.148	<.001	WesternEU	0.346	0.412	0.400
AE	0.295	0.116	0.011				

Table 2: Analysis of the Panitumumab Data

Table: Predicted survival functions for the panitumumab data:ITT: intent-to-treat, IPE: Branson and Whitehead (2002), Shao Cox: Shao et al. (2005), TM: proposed method, BSC: best supportive care alone, P+BSC: panitumumab plus best supportive care.

Time	ITT		No treatment switching		IPE		Shao Cox		TM	
(Days)	BSC	P+BSC	BSC	P+BSC	BSC	P+BSC	BSC	P+BSC	BSC	P+BSC
93	0·750	0·793	0·303	0·793	0·722	0·783	0·678	0·798	0·548	0·801
190	0·511	0·533	0·081	0·533	0·454	0·551	0·341	0·536	0·171	0·555
334	0·266	0·260	0·020	0·260	0·201	0·298	0·097	0·258	0·025	0·282
1024	0·013	0·038	0·020	0·038	0·001	0·007	0·001	0·023	0·001	0·026
P-value	0.·577		$< 0·001$		0·520		0·002		$< 0·001$	

Estimated Marginal Survival Curves



Motivation

- **Motivation 1**: In the panitumumab study, how do we quantify the overall treatment effect via a hazard ratio? We need to do this by estimating the marginal survival functions.
- **Motivation 2**: Motivation 1 implies that we can handle the case of crossing hazard functions
- **Motivation 3**: Suppose that different models are used for the same dataset – how do we compare the results from different models? For example, Abadi et al. (2012) used the Cox PH model, the accelerated failure time model, the generalized Gamma regression model, and the log-logistic regression for the same breast cancer dataset. Since the estimates obtained from these various models are not directly comparable, the authors had to rely completely on p-values for comparisons of the treatment effects.

Proposed Method

- What are known: $h_1(t)$ and $h_0(t)$ – the hazard functions in the experimental arm and control arm, respectively. Or equivalently, $S_1(t)$ and $S_0(t)$ – the marginal survival functions for the experimental and control arm.
- A general methodology for quantifying the average of time-varying hazard ratios is given by the following family of transformations,

$$\theta \equiv G^{-1}\left\{\int G\left(\frac{h_1(t)}{h_0(t)}\right)\Omega(t)dt\right\}, \tag{1}$$

  where $G(\cdot)$ is a strictly increasing transformation and $\Omega(t)$ is a weight function so that $\int \Omega(t)dt = 1$.
- In particular, we allow $\Omega(t)$ to depend on the underlying survival functions $S_1(t)$ and $S_0(t)$, and therefore, we write $\Omega(t)$ as $\Omega(t; S_0(t), S_1(t))$.

Proposed Method

- An interesting one-parameter transformation family is

$$G(x; a) = \{1 - (a + x)^{-a}\}/a,$$

where a is an unknown parameter.

- Since $\Omega(t)$ is a weight function, the transformation family can then be rewritten as

$$\theta_a = \left[\int \left\{ a + \frac{h_1(t)}{h_0(t)} \right\}^{-a} \Omega(t; S_0(t), S_1(t)) dt \right]^{-1/a} - a.$$

Proposed Method

- When a = -1, $G(x;-1) = x - 2$ yields the identity transformation

$$\theta_{iden} = \int \frac{h_1(t)}{h_0(t)} \Omega(t; S_0(t), S_1(t)) dt; \tag{2}$$

- When a = 0, $G(x;0) = \log(x)$ yields the logarithmic transformation

$$\theta_{log} = \exp\left[\int \log\left\{\frac{h_1(t)}{h_0(t)}\right\} \Omega(t; S_0(t), S_1(t)) dt\right]; \tag{3}$$

- When a = 1, $G(x;1) = x/(1+x)$ yields the ratio transformation

$$\theta_{ratio} = \frac{\int \{h_1(t)/h(t)\} \Omega\{t; S_0(t), S_1(t)\} dt}{\int \{h_0(t)/h(t)\} \Omega\{t; S_0(t), S_1(t)\} dt}, \tag{4}$$

where $h(t) = h_0(t) + h_1(t)$.

Proposed Method

- Under the proportional hazards assumption, i.e., when $h_1(t) = h_0(t)e^{\beta}$, $\theta_a = \theta_{iden} = \theta_{log} = \theta_{ratio} = e^{\beta}$.
- The identity transformation with $a = -1$ yields what Schemper et al. (2009) called the simple average hazard ratio.
- The logarithmic transformation with $a = 0$ yields the geometric average hazard ratio.
- The ratio transformation with $a = 1$, yields the average hazard ratio, which was originally defined in Kalbfleisch and Prentice (1981).
- The logarithmic and ratio transformations are the only two transformations among the one-parameter transformation family in (2) that are symmetric in $h_0(t)$ and $h_1(t)$. In other words, $\theta_a(h_0/h_1) = \{\theta_a(h_1/h_0)\}^{-1}$ when $a = 0$ or $a = 1$.

Proposed Method

- In the case when $-\log \hat{S}_k(t)$ is obtained nonparametrically using the Breslow estimator, we obtain

$$\hat{h}_k(t) = \int \frac{\sum_{i=1}^{n_k} K_{a_n}(s-t)dN_{ik}(s)}{\sum_{i=1}^{n_k} Y_{ik}(s)}, \quad k = 0, 1,$$

  where $N_{ik}(t)$ and $Y_{ik}(t)$ denote the observed counting process and at-risk process for subject i in treatment arm k, respectively.
- If marginal survival functions are available, as in our motivating panitumumab dataset, we can estimate $h_1(t)$ and $h_0(t)$ using the kernel estimates

$$\hat{h}_k(t) = -\int K_{a_n}(t-s)d\log \hat{S}_k(s), k = 0, 1,$$

  where $K_{a_n}(x) = a_n^{-1}K(x/a_n)$ for some non-negative and symmetric kernel function $K(x)$ with a_n being a bandwidth.

Proposed Method

- We also estimate $\Omega(t; S_0(t), S_1(t))$ using $\Omega(t; \hat{S}_0(t), \hat{S}_1(t))$.
- We obtain the estimator for general transformation $G(\cdot)$ as

$$\hat{\theta}_G = G^{-1}\left[\int G\left\{\frac{\hat{h}_1(t)}{\hat{h}_0(t)}\right\}\Omega\{t, \hat{S}_0(t), \hat{S}_1(t)\}dt\right]. \tag{5}$$

- Choice of the weight function $\Omega(t)$ was intensively studied in the literature. We pre-specify it to be proportional to $\{S_0(t)S_1(t)\}^{1/2}$.

Testing the Null Hypothesis of Identical Hazards

- We wish to test

$$H_0 : \text{identical hazards},$$

  which implies $\theta_a = 1$.
- We take $a \in [0, 1]$ since the test statistic has attractive large sample properties for $a \in [0, 1]$.
- Estimates with $a < 0$ tend to be numerically unstable and estimates with $a > 1$ impose large weights on local regions.

Testing the Null Hypothesis of Identical Hazards

- We define the estimate based on the maximum departure from the null as $\hat{\theta}_{sup} = \hat{\theta}_{\tilde{a}}$, where

$$\tilde{a} = \text{argsup}_{a \in [0,1]} \{ a : \mid \hat{\theta}_a - 1 \mid \}, \tag{6}$$

and

$$\hat{\theta}_a = \left[\int \left\{ a + \frac{\hat{h}_1(t)}{\hat{h}_0(t)} \right\}^{-a} \Omega(t; \hat{S}_0(t), \hat{S}_1(t)) dt \right]^{-1/a} - a. \tag{7}$$

- The test statistic based on the estimate with maximum departure from the null is a Kolmogorov−Smirnov type test statistic, which is given by

$$T_{sup} = \text{sup}_{a \in [0,1]} \mid \hat{\theta}_a - 1 \mid . \tag{8}$$

Asymptotic Results

Theorem

Under certain regularity conditions, $\hat{\theta}_G$ is consistent and $\sqrt{n}(\hat{\theta}_G - \theta_G)$ converges in distribution to a mean-zero normal distribution $N(0, \sigma^2)$.

Theorem

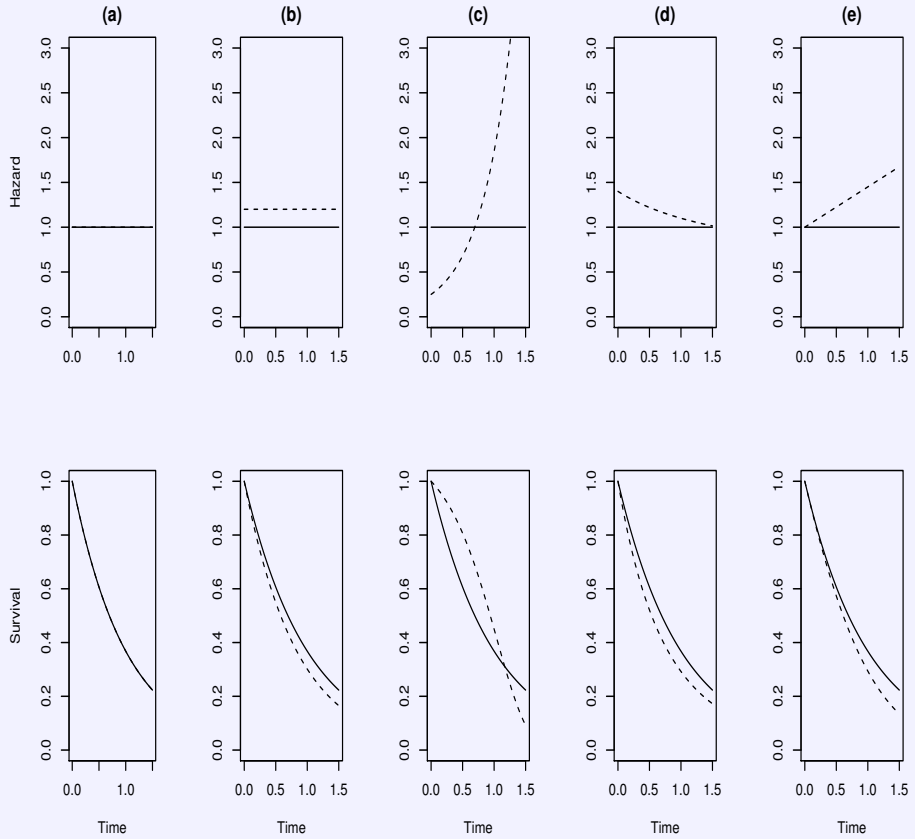
Under certain regularity conditions, for the transformation family with $a \in [-1, 1]$, $\hat{\theta}_a$ achieves its maximum local power at $a = 1$, the ratio transformation, when the weight function $\Omega(t)$ is independent of $S_0(t)$ and $S_1(t)$.

Theorem

Under certain regularity conditions and the null hypothesis of equal hazards, $\sqrt{n}T_{sup}$ converges in distribution to $sup_{a\in[0,1]} \mid \mathcal{G}_a \mid$, where $\mathcal{G}_a$ is a Gaussian process with mean 0.

Outline

Survival and Hazard Functions of Two Treatments with Five Typical Treatment Effects



- The hazard function of the control arm is defined by a constant hazard $h_0(t) = 1$ in each column.
- The hazard functions for the treatment arm are characterized by (a) $h_1(t) = h_0(t)$ as identical hazards; (b) $h_1(t) = 1.2h_0(t)$ as proportional hazards; (c) $h_1(t) = 0.25\exp(2t)h_0(t)$ as crossing hazards; (d) $h_1(t) = \{0.5 + 0.9/(1 + 0.5t)\}h_0(t)$ as converging hazards; (e) $h_1(t) = (1 + 0.45t)h_0(t)$ as diverging hazards.
- Censoring rates are 24%, 21%, 18.6%, 21.3%, and 19.6%, respectively, for (a)-(e).
- The bandwidth $a_n = 0.9\min(\hat{\sigma}, IQR/1.34)n^{-1/5}$, where $\hat{\sigma}$ and IQR are the standard deviation and the inter-quartile range of the variable in the kernel estimation, respectively.

Simulation Study I

- The 95% confidence intervals of the weighted hazard ratio estimates were based on normal approximations according to theoretical results we derived.
- For $\hat{\theta}_{sup}$, the normal approximation is not valid and, therefore, we constructed the 95% confidence interval using the 2.5% and 97.5% quantiles of the bootstrap samples.
- The simulation results are based on 1000 replicates.

Simulation Study I

Simulation Study I with $\Omega(t) \propto \{S_0(t)S_1(t)\}^{1/2}$ and sample size $n = 800$.

Method	TRUE	BIAS ($\times$1000)	SD ($\times$100)	ESE ($\times$100)	CP %	Type I ($\times$100)
		(a) Identical Hazards; CR=24%				
Cox	1.00	5	8.2	8.2	95	5
Identity ($a = -1$)	1.00	15	8.4	8.7	96	4
Logarithm ($a = 0$)	1.00	5	8.3	8.5	96	4
Ratio ($a = 1$)	1.00	5	8.3	8.4	96	4
$\hat{\theta}_{sup}$	1.00	5	8.3	8.5	95	5

Simulation Study I

Simulation Study I with $\Omega(t) \propto \{S_0(t)S_1(t)\}^{1/2}$ and sample size $n = 800$.

Method	TRUE	BIAS ($\times$1000)	SD ($\times$100)	ESE ($\times$100)	CP% %	Power ($\times$100)
		(b) Proportional Hazards; CR=21%				
Cox	1.20	6	9.7	9.6	94	58
Identity ($a = -1$)	1.20	19	10.0	10.4	96	57
Logarithm ($a = 0$)	1.20	7	9.9	10.1	96	52
Ratio ($a = 1$)	1.20	6	9.8	10.0	96	52
$\hat{\theta}_{sup}$	1.20	7	9.9	10.1	95	62
		(c) Crossing Hazards; CR=18.6%				
Cox	1.00	6	8.1	8.0	95	5
Identity ($a = -1$)	1.05	33	10.7	11.1	97	6
Logarithm ($a = 0$)	0.76	29	7.2	7.0	93	82
Ratio ($a = 1$)	0.77	25	6.4	6.2	94	86
$\hat{\theta}_{sup}$	0.76	28	7.1	6.8	92	81
		d) Converging Hazards; CR=21.3%				
Cox	1.21	21	9.9	9.8	95	66
Identity ($a = -1$)	1.22	15	10.2	10.5	96	63
Logarithm ($a = 0$)	1.22	3	10.1	10.4	96	57
Ratio ($a = 1$)	1.22	3	10.0	10.2	96	58
$\hat{\theta}_{sup}$	1.22	4	10.1	10.4	95	65
		(e) Diverging Hazards; CR=19.6%				
Cox	1.24	5	10.0	9.9	94	72
Identity ($a = -1$)	1.25	24	10.4	10.9	96	73
Logarithm ($a = 0$)	1.23	12	10.0	10.3	96	68
Ratio ($a = 1$)	1.23	11	9.9	10.2	96	69
$\hat{\theta}_{sup}$	1.23	13	10.0	10.4	95	75

Simulation Study I with $\Omega(t) \propto \{S_0(t)S_1(t)\}^{1/2}$ and sample size $n = 800$.

Method	TRUE	BIAS ($\times$1000)	SD ($\times$100)	ESE ($\times$100)	CP% %	Power ($\times$100)
		(b) Proportional Hazards; CR=21%				
Cox	1.20	6	9.7	9.6	94	58
Identity ($a = -1$)	1.20	19	10.0	10.4	96	57
Logarithm ($a = 0$)	1.20	7	9.9	10.1	96	52
Ratio ($a = 1$)	1.20	6	9.8	10.0	96	52
$\hat{\theta}_{sup}$	1.20	7	9.9	10.1	95	62
		(c) Crossing Hazards; CR=18.6%				
Cox	1.00	6	8.1	8.0	95	5
Identity ($a = -1$)	1.05	33	10.7	11.1	97	6
Logarithm ($a = 0$)	0.76	29	7.2	7.0	93	82
Ratio ($a = 1$)	0.77	25	6.4	6.2	94	86
$\hat{\theta}_{sup}$	0.76	28	7.1	6.8	92	81
		d) Converging Hazards; CR=21.3%				
Cox	1.21	21	9.9	9.8	95	66
Identity ($a = -1$)	1.22	15	10.2	10.5	96	63
Logarithm ($a = 0$)	1.22	3	10.1	10.4	96	57
Ratio ($a = 1$)	1.22	3	10.0	10.2	96	58
$\hat{\theta}_{sup}$	1.22	4	10.1	10.4	95	65
		(e) Diverging Hazards; CR=19.6%				
Cox	1.24	5	10.0	9.9	94	72
Identity ($a = -1$)	1.25	24	10.4	10.9	96	73
Logarithm ($a = 0$)	1.23	12	10.0	10.3	96	68
Ratio ($a = 1$)	1.23	11	9.9	10.2	96	69
$\hat{\theta}_{sup}$	1.23	13	10.0	10.4	95	75

- When the alternatives are proportional hazards, converging hazards or diverging hazards, the ratio and logarithmic transformations yield little loss of power compared to the identity transformation and the Cox model.

Simulation Study I with $\Omega(t) \propto \{S_0(t)S_1(t)\}^{1/2}$ and sample size $n = 800$.

Method	TRUE	BIAS ($\times$1000)	SD ($\times$100)	ESE ($\times$100)	CP% %	Power ($\times$100)
		(b) Proportional Hazards; CR=21%				
Cox	1.20	6	9.7	9.6	94	58
Identity ($a = -1$)	1.20	19	10.0	10.4	96	57
Logarithm ($a = 0$)	1.20	7	9.9	10.1	96	52
Ratio ($a = 1$)	1.20	6	9.8	10.0	96	52
$\hat{\theta}_{sup}$	1.20	7	9.9	10.1	95	62
		(c) Crossing Hazards; CR=18.6%				
Cox	1.00	6	8.1	8.0	95	5
Identity ($a = -1$)	1.05	33	10.7	11.1	97	6
Logarithm ($a = 0$)	0.76	29	7.2	7.0	93	82
Ratio ($a = 1$)	0.77	25	6.4	6.2	94	86
$\hat{\theta}_{sup}$	0.76	28	7.1	6.8	92	81
		d) Converging Hazards; CR=21.3%				
Cox	1.21	21	9.9	9.8	95	66
Identity ($a = -1$)	1.22	15	10.2	10.5	96	63
Logarithm ($a = 0$)	1.22	3	10.1	10.4	96	57
Ratio ($a = 1$)	1.22	3	10.0	10.2	96	58
$\hat{\theta}_{sup}$	1.22	4	10.1	10.4	95	65
		(e) Diverging Hazards; CR=19.6%				
Cox	1.24	5	10.0	9.9	94	72
Identity ($a = -1$)	1.25	24	10.4	10.9	96	73
Logarithm ($a = 0$)	1.23	12	10.0	10.3	96	68
Ratio ($a = 1$)	1.23	11	9.9	10.2	96	69
$\hat{\theta}_{sup}$	1.23	13	10.0	10.4	95	75

- The test based on $\hat{\theta}_{sup}$ is more powerful in proportional hazards, converging hazards or diverging hazards, with a similar magnitude of power loss in crossing hazards comparing to the ratio and logarithmic transformations.

Simulation Study I with $\Omega(t) \propto \{S_0(t)S_1(t)\}^{1/2}$ and sample size $n = 800$.

Method	TRUE	BIAS ($\times$1000)	SD ($\times$100)	ESE ($\times$100)	CP% %	Power ($\times$100)
		(b) Proportional Hazards; CR=21%				
Cox	1.20	6	9.7	9.6	94	58
Identity ($a = -1$)	1.20	19	10.0	10.4	96	57
Logarithm ($a = 0$)	1.20	7	9.9	10.1	96	52
Ratio ($a = 1$)	1.20	6	9.8	10.0	96	52
$\hat{\theta}_{sup}$	1.20	7	9.9	10.1	95	62
		(c) Crossing Hazards; CR=18.6%				
Cox	1.00	6	8.1	8.0	95	5
Identity ($a = -1$)	1.05	33	10.7	11.1	97	6
Logarithm ($a = 0$)	0.76	29	7.2	7.0	93	82
Ratio ($a = 1$)	0.77	25	6.4	6.2	94	86
$\hat{\theta}_{sup}$	0.76	28	7.1	6.8	92	81
		d) Converging Hazards; CR=21.3%				
Cox	1.21	21	9.9	9.8	95	66
Identity ($a = -1$)	1.22	15	10.2	10.5	96	63
Logarithm ($a = 0$)	1.22	3	10.1	10.4	96	57
Ratio ($a = 1$)	1.22	3	10.0	10.2	96	58
$\hat{\theta}_{sup}$	1.22	4	10.1	10.4	95	65
		(e) Diverging Hazards; CR=19.6%				
Cox	1.24	5	10.0	9.9	94	72
Identity ($a = -1$)	1.25	24	10.4	10.9	96	73
Logarithm ($a = 0$)	1.23	12	10.0	10.3	96	68
Ratio ($a = 1$)	1.23	11	9.9	10.2	96	69
$\hat{\theta}_{sup}$	1.23	13	10.0	10.4	95	75

- We were surprised to observe superior power of $\hat{\theta}_{sup}$ over the standard Cox model under proportional hazards.

Remarks on Simulation Study I

- When the alternative is crossing hazards, the ratio and logarithmic transformations yield substantial power gains compared to the identity transformation and the Cox model.
- When the alternatives are proportional hazards, converging hazards or diverging hazards, the ratio and logarithmic transformations yield little loss of power compared to the identity transformation and the Cox model.
- The test based on $\hat{\theta}_{sup}$ is more powerful in proportional hazards, converging hazards or diverging hazards, with a similar magnitude of power loss in crossing hazards comparing to the ratio and logarithmic transformations.
- We were surprised to observe superior power of $\hat{\theta}_{sup}$ over the standard Cox model under proportional hazards.

Simulation Study II

Simulation Study II with $\Omega(t) \propto 1$ and sample size $n = 400$

Method	TRUE	BIAS ($\times 1000$)	SD ($\times 100$)	ESE ($\times 100$)	CP% %	Type I ($\times 100$)
			(a) Identical Hazards; CR=24%			
Cox	1.00	8	11.3	11.6	96	4
Identity ($a = -1$)	1.00	31	13.5	14.9	97	3
Logarithm ($a = 0$)	1.00	8	12.9	13.7	96	4
Ratio ($a = 1$)	1.00	7	12.7	13.3	95	5
$\hat{\theta}_{sup}$	1.23	8	12.9	13.7	95	5

Simulation Study II with $\Omega(t) \propto 1$ and sample size $n = 400$

Method	TRUE	BIAS ($\times 1000$)	SD ($\times 100$)	ESE ($\times 100$)	CP% %	Power ($\times 100$)
		(b) Proportional Hazards; CR=21%				
Cox	1.20	10	13.4	13.7	96	27
Identity ($a = -1$)	1.20	37	16.2	18.3	98	12
Logarithm ($a = 0$)	1.20	9	15.5	16.7	96	14
Ratio ($a = 1$)	1.20	6	15.2	16.2	96	15
$\hat{\theta}_{sup}$	1.20	9	15.5	16.7	96	25
		(c) Crossing Hazards; CR=18.6%				
Cox	1.00	13	11.5	11.4	95	5
Identity ($a = -1$)	1.59	79	31.4	36.4	97	44
Logarithm ($a = 0$)	1.12	48	14.3	15.2	97	12
Ratio ($a = 1$)	1.10	38	11.5	12.0	97	15
$\hat{\theta}_{sup}$	1.12	50	14.3	15.2	95	22
		(d) Converging Hazards; CR=21.3%				
Cox	1.21	26	13.7	14.0	96	34
Identity ($a = -1$)	1.17	36	15.6	17.4	97	10
Logarithm ($a = 0$)	1.17	9	15.5	16.5	96	9
Ratio ($a = 1$)	1.17	8	15.3	16.0	95	10
$\hat{\theta}_{sup}$	1.17	9	15.5	16.5	95	18
		(e) Diverging Hazards; CR=19.6%				
Cox	1.24	9	13.7	14.0	95	39
Identity ($a = -1$)	1.34	46	19.2	21.9	98	32
Logarithm ($a = 0$)	1.32	13	16.7	18.2	96	41
Ratio ($a = 1$)	1.32	8	16.0	17.4	96	44
$\hat{\theta}_{sup}$	1.32	13	16.6	18.2	96	58

- The choice of weight functions has a great influence on the estimates as studied in Lininger et al. (1979), Pepe and Fleming (1989), Shen and Fleming (1997).
- Comparisons within Cox model and identity/logarithmic/ratio transformation are not as clear as the weight $\propto \{S_0(t)S_1(t)\}^{1/2}$.

Simulation Study II with $\Omega(t) \propto 1$ and sample size $n = 400$

Method	TRUE	BIAS ($\times 1000$)	SD ($\times 100$)	ESE ($\times 100$)	CP% %	Power ($\times 100$)
		(b) Proportional Hazards; CR=21%				
Cox	1.20	10	13.4	13.7	96	27
Identity ($a = -1$)	1.20	37	16.2	18.3	98	12
Logarithm ($a = 0$)	1.20	9	15.5	16.7	96	14
Ratio ($a = 1$)	1.20	6	15.2	16.2	96	15
$\hat{\theta}_{sup}$	1.20	9	15.5	16.7	96	25
		(c) Crossing Hazards; CR=18.6%				
Cox	1.00	13	11.5	11.4	95	5
Identity ($a = -1$)	1.59	79	31.4	36.4	97	44
Logarithm ($a = 0$)	1.12	48	14.3	15.2	97	12
Ratio ($a = 1$)	1.10	38	11.5	12.0	97	15
$\hat{\theta}_{sup}$	1.12	50	14.3	15.2	95	22
		(d) Converging Hazards; CR=21.3%				
Cox	1.21	26	13.7	14.0	96	34
Identity ($a = -1$)	1.17	36	15.6	17.4	97	10
Logarithm ($a = 0$)	1.17	9	15.5	16.5	96	9
Ratio ($a = 1$)	1.17	8	15.3	16.0	95	10
$\hat{\theta}_{sup}$	1.17	9	15.5	16.5	95	18
		(e) Diverging Hazards; CR=19.6%				
Cox	1.24	9	13.7	14.0	95	39
Identity ($a = -1$)	1.34	46	19.2	21.9	98	32
Logarithm ($a = 0$)	1.32	13	16.7	18.2	96	41
Ratio ($a = 1$)	1.32	8	16.0	17.4	96	44
$\hat{\theta}_{sup}$	1.32	13	16.6	18.2	96	58

- When comparing $\hat{\theta}_{sup}$ to the estimates based on the logarithmic and ratio transformations, $\hat{\theta}_{sup}$ outperforms the other two estimates in terms of power in all settings.

Outline

Panitumumab Data Revisited

The estimates of $\theta(h_0/h_1; a)$ with the outcome of overall survival and weight function proportional to $\{S_0(t)S_1(t)\}^{1/2}$.

Transformation	Estimate	SD	95% CI	P-value
Logarithm	2.90	0.84	(1.48, 4.73)	0.008
Ratio	2.91	0.81	(1.61, 4.69)	$<$0.001
$\hat{\theta}_{sup}$	2.92	0.84	(1.61, 4.80)	$<$0.001

- The estimates, standard errors, 95% confidence intervals, and p-values based on 1000 bootstrap samples
- Note that the bootstrap sampling procedure was conducted on the original dataset and the marginal survival curves were constructed for each bootstrap sample. Therefore, the reported 95% confidence intervals and p-values have incorporated the variation associated with the estimation of the marginal survival functions.

Panitumumab Data Revisited

- Due to the complexity of the data, survival models such as the Cox proportional hazards model or weighted Cox proportional hazards model are no longer valid, and hence, were not included in Table.
- As shown in the table, the hazard ratio estimates based on different transformations are very stable, ranging from 2.90 to 2.92, with highly significant p-values. Judging from the p-values, the ratio-transformed estimate and the maximum departure-based estimate are a little more powerful than the logarithm-transformed estimate, although the difference may be of limited practical relevance in this study.

Outline

Concluding Remarks

- Our numerical studies results suggested that when the weight function is proportional to $\{S_0(t)S_1(t)\}^{1/2}$, the ratio transformation tended to provide a larger power in distinguishing between the two treatment arms compared to the identity and logarithmic transformations.
- We also demonstrated that the ratio transformation achieves the maximum local power within the transformation family when the weight function is independent of $S_0(t)$ and $S_1(t)$.
- The simulation studies show that when the hazard functions of the two groups either converge or diverge, $\hat{\theta}_{sup}$ is more powerful than the test statistic based on the individual transformations, with a similar magnitude of power loss when the hazards cross, demonstrating the importance of correctly identifying the shape of the time-varying hazard ratio function for estimation.

Concluding Remarks

- Choosing $G(x)$ should not only depend on the resulting statistical power under the alternative, but it also needs to yield a good interpretation of the parameters. For reasons of interpretability, we recommend using the ratio-transformed estimator when estimation is of primary interest, and to use the Kolmogorov-Smirnov type test statistic T_{sup} when hypothesis testing is of primary interest.
- Although we considered a class of transformations to quantify the average hazard ratios, the same concept can be generalized to quantify many other time-varying comparison measures. These include time-varying intensities for recurrent events, time varying treatment effects over time, and time-sensitive diagnostic measures.