A GENERAL FRAMEWORK FOR SEQUENTIAL AND ADAPTIVE METHODS IN SURVIVAL STUDIES

By Xiaolong Luo, Gongjun Xu and Zhiliang Ying*

Celgene Corporation, Columbia University and Columbia University

Adaptive treatment allocation schemes based on interim responses have generated a great deal of recent interest in clinical trials and other follow-up studies. An important application of such schemes is in survival studies, where the response variable of interest is time to the occurrence of a certain event. Due to possible dependency structures inherited from the enrollment and allocation schemes, existing approaches to survival models, including those that handle staggered entry, cannot be applied directly. This paper develops a new general framework with its theoretical foundation for handling such adaptive designs. The new approach is based on marked point processes and differs from existing approaches in that it considers entry and calender times rather than survival and calender times. Large sample properties, which are essential for statistical inference, are established. Special attention is given to the Cox model and related score processes. Applications to adaptive and sequential designs are discussed.

1. Introduction. Sequential and adaptive methods play important roles in the design and analysis of long-term clinical studies. Pocock (1977), O'Brien and Fleming (1979), and Lan and DeMets (1983) proposed various boundaries that adjust for multiple testing and are motivated by applications to clinical trials; see also Jennison and Turnbull (2000). Zelen (1969), Wei and Durham (1978), and Wei (1978) proposed and studied outcome dependent treatment allocation schemes; see also Hu and Rosenberger (2006), Rosenberger and Sverdlov (2008), and Hu, Zhang and He (2009). Fisher (1998), Cui, Hung and Wang (1999), and Shen and Fisher (1999), on the other hand, proposed adaptive schemes under which sample sizes are re-estimated and adapted to interim analysis. Robins (1986) and Murphy and Bingham (2008) developed dynamic treatment regimes, in which treatment allocations are dynamically adapted to interim outcomes.

Many long-term clinical trials and epidemiological cohort studies involve survival endpoints; see Kalbfleisch and Prentice (2002) and Fleming and

*Research supported in part by the National Institutes of Health.

AMS 2000 subject classifications: Primary 62L10, 62E20.

Keywords and phrases: Sequential analysis, adaptive allocation, marked point process, proportional hazards regression, staggered entry, weak convergence.

Harrington (1991) for examples of such kind and standard statistical methods. For survival data, the log-rank test (Mantel and Haenszel, 1959) and proportional hazards regression (Cox, 1972) are the methods of choice. Counting processes and associated martingales may be used to derive desired theoretical properties; see Andersen, Borgan, Gill and Keiding (1993).

Sequential and adaptive methods for survival data require simultaneous consideration of both calendar and survival times. Sellke and Siegmund (1983) and Slud (1984) established Brownian approximation to the score process calculated over the diagonal line, i.e. when calendar time meets survival time. A Gaussian random field approximation to the two-dimensional score process in the case of two-sample comparison was established by Gu and Lai (1991). More general results about Gaussian random field approximation to the two-dimensional score process under the Cox proportional hazards regression can be found in Bilias, Gu and Ying (1997), who made use of modern empirical process theory to derive certain key results, bypassing martingale formulation to handle asymptotic analysis.

To incorporate both group sequential analysis and adaptive outcome-dependent treatment allocation, one needs to consider a score process, to which the existing martingale approach or the empirical process theory is not applicable. That empirical process theory may not be applied is largely due to the outcome-dependent enrollment allocation, which results in study units that are not mutually independent.

This paper develops a new theoretic framework and techniques for the partial likelihood score process with simultaneous consideration of calendar and survival times and with entry times and treatment allocation possibly depending on preceding outcomes. The approach is based upon expressing the score process as a stochastic integral of a suitably defined marked point process. The use of marked point processes for survival data was introduced by Arjas and Haara (1984) and Arjas (1989). Related subsequent developments can be found in Feng (1999) and Martinussen and Scheike (2006).

The paper is organized as follows. Section 2 provides basic notation and an introduction to the marked point process framework under calender time. It also gives an illustrative example involving the Cox model. The corresponding functional central limit theorems in a general setting are presented in Section 3. An application to the Cox proportional hazards model with time-dependent covariates is given in Section 4, where convergence properties for the corresponding maximum partial likelihood estimator are also established. Some discussion and more applications are given in Section 5. Proofs for our main results are provided in Section 6.

2. Notation and marked point process framework for survival data.

2.1. *Marked point process framework.* Consider a follow up study with calendar time period $[0, \mathcal{T})$, $\mathcal{T} \le \infty$. Let U_i be the entry time for individual $i, i \ge 1$. For technical convenience, we assume throughout this paper that the U_i have no ties. Thus, without loss of generality, we assume $U_1 < U_2 < \cdots < U_i < \cdots$. Define the associated counting process for entry times

$$R_t = \sum_{i \ge 1} 1(U_i \le t). \tag{2.1}$$

For subject i, let T_i denote survival time (since entry) and C_i censoring time. Let $\tilde{T}_i = T_i \wedge C_i$ and $\Delta_i = 1(T_i \le C_i)$, indicating failure (1) or censoring (0). Thus, if $\Delta_i = 1(0)$, then individual i experiences failure (censoring) at calendar time $U_i + \tilde{T}_i$. Furthermore, there is a possibly time-dependent, d-dimensional covariate vector $Z_i = Z_i(\cdot)$, which may include the i-th individual's treatment assignment and certain relevant baseline characteristics. Here, for any $w \ge 0$, $Z_i(w)$ refers to the covariate value at calendar time $U_i + w$. As usual, we assume that Z_i, as a random variable, is non-informative for T_i and C_i (i.e. external time-dependent covariates as discussed in Kalbfleisch and Prentice, 2002). Informally speaking, the Z_i can be taken as a deterministic prior to information for T_i and C_i, and thus will be probabilistically independent of future event development and follow up, though it may depend on the historical information up to time U_i.

With the above notation, the entire underlying collection of random variables that may be observed with sufficiently long follow up is $\{U_i, Z_i, U_i + \tilde{T}_i, \Delta_i, i \ge 1\}$. The data may be visualized as a two dimensional plot of entry time U_i and event time $U_i + \tilde{T}_i$, with each point labeled by (Z_i, Δ_i). Since it is the event, not entry, time that is of interest here, we combine U_i with (Z_i, Δ_i) to form a point (U_i, Z_i, Δ_i) marking the event time $U_i + \tilde{T}_i$ in the mark space $\mathcal{X}$. The structure of $\mathcal{X}$ is built as follows. The first component is a real number in $[0, \mathcal{T})$. The third component is either 0 or 1. The second component Z_i is a function in $\mathcal{D}_z$, the set of all right continuous functions on $[0, \mathcal{T})$ with bounded variation, on which we use the Skorohod topology to define a measurable space. Thus, the mark space $\mathcal{X} = [0, \mathcal{T}) \times \mathcal{D}_z \times \{0, 1\}$ has a naturally induced product σ algebra.

For a given $t > 0$, define a random counting measure on $\mathcal{X}$

$$p_t(I \times E) \quad = \quad \sum_i 1\left(U_i + \tilde{T}_i \le t, U_i \in I, (Z_i, \Delta_i) \in E\right),$$

where $I \subset [0, \mathcal{T})$ and $E \subset \mathcal{D}_z \times \{0, 1\}$ are the corresponding Borel measurable subsets. Note that p_t is random since it depends on random variables U_i, $\tilde{T}_i$, Z_i and Δ_i, $i \geq 1$. From (2.1), we can rewrite

$$p_t(I \times E) \quad = \quad \int_I 1\left(u + \tilde{T}_u \leq t, (Z_u, \Delta_u) \in E\right) dR_u, \tag{2.2}$$

where, with a slight abuse of notation, $\tilde{T}_u$, Z_u, and Δ_u refer to $\tilde{T}_i$, Z_i, and Δ_i when $u = U_i$, which is well defined since the U_i are distinct for different i.

The random measure $p_t(I \times E)$ may be viewed as a trivariate function of t, I, and E. When I and E are fixed, $p_t(I \times E)$ is a non-decreasing function of t, thereby inducing a Lebesgue-Stieltjes measure on $[0, \mathcal{T})$. Since for fixed t, $p_t(I \times E)$ is already a measure on $\mathcal{X}$, we can combine the two measures together to get a joint measure on $[0, \mathcal{T}) \times \mathcal{X}$

$$\begin{aligned} p(ds\, du\, dz\, d\delta) &= \sum_{i \geq 1} 1(U_i + \tilde{T}_i \in ds, U_i \in du, Z_i \in dz, \Delta_i = \delta) \\ &= 1(u + \tilde{T}_u \in ds, Z_u \in dz, \Delta_u = \delta) dR_u. \end{aligned} \tag{2.3}$$

Note that the support of this measure is on $\{(u, s) : 0 \leq u \leq s < \mathcal{T}\}$.

The above random measure also provides a way to define information accumulation over calendar time t. This is done by introducing the following internal σ-filtration.

$$\begin{aligned} \mathcal{F}_t &= \sigma\Big\{p(A \times I \times E), Z_u, R_u : \\ &\qquad \forall u \leq t, A \subset [0, t], I \subset [0, t], E \subset \mathcal{D}_z \times \{0, 1\}\Big\}. \end{aligned}$$

A sub-σ-algebra of $\mathcal{F}_t$ that is of interest is defined by

$$\begin{aligned} \mathcal{F}_{t,\vartheta} &= \sigma\Big\{p(A \times I \times E), Z_u, R_u : \\ &\qquad \forall u \leq \vartheta \wedge t, A \times I \subset [0, t] \times [0, u], E \subset \mathcal{D}_z \times \{0, 1\}\Big\}. \end{aligned}$$

Intuitively, $\mathcal{F}_{t,\vartheta}$ represents covariate and event history up to time t for individuals that enrolled before time ϑ, where $0 < \vartheta \leq t$.

Without loss of generality, we shall assume throughout that R_t and Z_t are predictable with respect to $\{\mathcal{F}_t, t \geq 0\}$, which is standard in survival analysis. We also need the following condition.

Condition A. An individual's current survival probability does not depend on the future information of people who enrolled earlier than him/her in the sense that for any s and $t > s$,

$$P(\tilde{T}_i \in (s, s+ds], \Delta_i | \mathcal{F}_{U_i+t, U_i-}) = P(\tilde{T}_i \in (s, s+ds], \Delta_i | \mathcal{F}_{U_i+s, U_i-}). \tag{2.4}$$

Moreover, assume

$$P(\tilde{T}_i \in (s, s+ds], \Delta_i | \mathcal{F}_{U_i+s}) = P(\tilde{T}_i \in (s, s+ds], \Delta_i | \mathcal{F}_{U_i+s, U_i}). \tag{2.5}$$

REMARK 2.1. *Here* $\mathcal{F}_{U_i+s,U_i-}$ *represents covariates and event history up to calendar time* U_i+s *for subjects* $1, \cdots, i-1$; $\mathcal{F}_{U_i+s,U_i}$ *represents covariates and event history up to calendar time* $U_i + s$ *for subjects* $1, \cdots, i-1, i$.

As we mentioned earlier, $p_t(I \times E)$ is non-decreasing in t. Therefore, the Dood-Meyer decomposition (Jacod and Shiryaev, 2003, page 66) implies the existence of a compensator. The following lemma provides a general way to obtain such a compensator and the corresponding martingale properties.

LEMMA 2.2. *For any Borel set* $I \subset [0, \mathcal{T})$ *and* $E \subset \mathcal{D}_z \times \{0, 1\}$, *there exists a predictable compensator* $q(ds\, du\, dz\, d\delta)$ *for* $p(ds\, du\, dz\, d\delta)$, *such that*

$$M_t(I \times E) \triangleq p_t(I \times E) - \int_0^t \int_{I \times E} q(ds\, du\, dz\, d\delta) \tag{2.6}$$

is a $\{\mathcal{F}_t, t \geq 0\}$ *martingale, where* $q(dt\, du\, dz\, d\delta) = E(p(dt\, du\, dz\, d\delta)|\mathcal{F}_{t-})$. *Moreover, under Condition A, for fixed* t,

$$M_{t,\vartheta}(E) \triangleq p_t([0, \vartheta] \times E) - \int_0^t \int_0^{\vartheta \wedge s} \int_E q(ds\, du\, dz\, d\delta), \tag{2.7}$$

as a process in ϑ, *is a* $\{\mathcal{F}_{t,\vartheta}, 0 \leq \vartheta \leq t\}$ *martingale.*

REMARK 2.3. *Here we define the basic martingale process in calendar time, in contrast to the usual approach of defining the basic martingale process in survival time. The calendar time-based approach is more natural for sequential analysis since interim analyses are conducted along calendar time. In addition, we use the entry time as the second time dimension. This is also natural since entry time indicates sample accumulation.*

From (2.6), $M_t(\cdot)$ defines a combined random measure $dM_s = p(ds\, du\, dz\, d\delta) - q(ds\, du\, dz\, d\delta)$ on the calendar time and mark space $[0, \mathcal{T}) \times \mathcal{X}$. More formally, by a random measure on $[0, \mathcal{T}) \times \mathcal{X}$, we mean a kernel mapping from the event space Ω to $[0, \mathcal{T}) \times \mathcal{X}$ (Last and Brandt, 1995). At point

$(s,u,z,\delta) \in [0,\mathcal{T}) \times \mathcal{X}$, let $f_{s,u,z,\delta}(t)$ be an $\mathcal{F}_t$ measurable random variable indexed by (s,u,z,δ); see (2.14) for the Cox model as an example. Its integral with respect to dM_s can be expressed as

$$\begin{aligned}\int_0^t \int_{I\times E} f_{s,u,z,\delta}(t)dM_s &= \sum_{\substack{i,U_i+\tilde{T}_i\le t,\\ U_i\in I,(Z_i,\Delta_i)\in E}} f_{U_i+\tilde{T}_i,U_i,Z_i,\Delta_i}(t)\\ &\quad - \int_0^t \int_{I\times E} f_{s,u,z,\delta}(t)q(ds\,du\,dz\,d\delta).\end{aligned} \tag{2.8}$$

When $f_{s,u,z,\delta}(t)$ is $\mathcal{F}_s$ predictable, results for martingale integration may be used; see Kallianpur and Xiong (1995, Chapter 3) and Jacod and Shiryaev (2003, Chapter II). In particular, for $\mathcal{F}_s$ predictable $f_{s,u,z,\delta}(t)$ with

$$E\int_0^{\mathcal{T}} \int_{\mathcal{X}} f^2_{s,u,z,\delta}(t)p(ds\,du\,dz\,d\delta) < \infty,$$

the above integral

$$M_t^f(I\times E) \triangleq \int_0^t \int_{I\times E} f_{s,u,z,\delta}(t)dM_s \tag{2.9}$$

is a square integrable $\{\mathcal{F}_t, t\ge 0\}$ martingale with predictable variation process

$$\langle M^f(I\times E)\rangle(t) = \int_0^t \int_{I\times E} f^2_{s,u,z,\delta}(t)q(ds\,du\,dz\,d\delta), \tag{2.10}$$

which is useful for variance estimation.

In general, the predictability assumption may not always be satisfied. In those cases, we will use dM_s as a measure for sample path-wise integration, which is well defined in (2.8).

2.2. *Cox proportional hazards regression model.* We illustrate the above construction through the Cox (1972) proportional hazards regression model with a dependent/independent enrollment process. For simplicity, we take Z_i to be one-dimensional. For survival time T_i, the Cox model specifies

$$P(T_i > w|\mathcal{F}_{U_i+w}) = \exp\left\{-\int_0^w \exp\{\beta Z_i(s)\}\lambda_0(s)ds\right\}, w>0, \tag{2.11}$$

where $\lambda_0(.)$ is the baseline hazard function and β is the regression parameter. In addition, we use $\lambda_{i,c}(.)$ to denote the hazard function for the censoring time C_i.

By Lemma 2.2, we can write the compensator for $p(ds\,du\,dz\,d\delta)$ as

$$q(ds\,du\,dz\,d\delta)$$
$$= \begin{cases} 1(\tilde{T}_u \geq s-u, Z_u \in dz)\exp\{\beta Z_u(s-u)\}\lambda_0(s-u)dR_u ds & s \geq u, \delta = 1; \\ 1(\tilde{T}_u \geq s-u, Z_u \in dz)\lambda_{u,c}(s-u)dR_u ds & s \geq u, \delta = 0; \\ 0 & \text{otherwise.} \end{cases}$$

For each $k = 0, 1, 2$ and any $\vartheta > 0, w > 0, \vartheta + w < \mathcal{T}$, let

$$\Gamma_k(\beta; \vartheta, w) = \sum_{U_i \leq \vartheta} Z_i^k(w)\exp(\beta Z_i(w))1(\tilde{T}_i \geq w). \tag{2.12}$$

We can express the log partial likelihood $l(\beta; t)$ for β as

$$\int_0^t \int_0^s \int_{\mathcal{D}_z \times \{0,1\}} \Big(\beta Z_u(s-u) - \log[\Gamma_0(\beta; t-(s-u), s-u)]\Big) 1(\delta = 1) p(ds\,du\,dz\,d\delta); \tag{2.13}$$

see equation (1) in Sellke and Siegmund (1983). The score process can then be written as

$$\begin{aligned} U(\beta; t) &= \int_0^t \int_0^s \int_{\mathcal{D}_z \times \{0,1\}} [Z_u(s-u) - \bar{Z}(\beta; t, s-u)]1(\delta = 1)p(dsdudzd\delta) \\ &= \int_0^t \int_0^s \int_{\mathcal{D}_z \times \{1\}} f_{s,u,z,\delta}(t)dM_s, \end{aligned}$$

where $dM_s = p(ds\,du\,dz\,d\delta) - q(ds\,du\,dz\,d\delta)$ and $f_{s,u,z,\delta}(t) = z(s-u) - \bar{Z}(\beta; t, s-u)$, in which $z(\cdot)$ is the index function z in $f_{s,u,z,\delta}(t)$ and

$$\bar{Z}(\beta; t, w) = \frac{\Gamma_1(\beta; t-w, w)}{\Gamma_0(\beta; t-w, w)}.$$

More generally, we can define a two-parameter score process with respect to calendar time t and entry time ϑ as

$$U(\beta; t, \vartheta) = \int_0^t \int_0^{s \wedge \vartheta} \int_{\mathcal{D}_z \times \{1\}} f_{s,u,z,\delta}(t)dM_s. \tag{2.14}$$

Note that $U(\beta; t, t) = U(\beta; t)$.

As we mentioned in Remark 2.3, $U(\beta; t)$ is an integral along calendar time instead of survival time. By Lemma 2.2, we can use the martingale central limit theorem to obtain convergence properties for $U(\beta; t)$ in t; see Sections

3 and 4 for details. Through this framework, enrollment and covariates history is expressed by the filtration $\mathcal{F}_t$. As a result, martingale structure still holds under response and covariates dependent allocation scheme, which is desirable for adaptive methods in clinical trials (cf. Hu and Rosenberger, 2006).

On the other hand, the usual survival time-based approach results in the following score process

$$U_n(\beta; t, t) = \sum_{U_i \le t} \int_0^t [Z_i(s) - \bar{Z}(\beta; t, s)] N_i(t, ds), \tag{2.15}$$

where $N_i(t, s) = \Delta_i 1(\tilde{T}_i \le s \wedge (t - U_i)^+)$; see Bilias, et al., (1997). The underlying martingale processes are

$$m_i(t, s) = N_i(t, s) - \int_0^s 1(\tilde{T}_i \wedge (t - U_i)^+ \ge w) \exp(\beta Z_i(w)) \lambda_0(w) dw \tag{2.16}$$

with filtration $\{\mathcal{F}_t(s), s > 0\}$ containing all information up to survival time s and calendar time t for all subjects enrolled before t.

Under (2.15) and (2.16), if enrollment process follows a response and/or covariate adaptive randomization procedure, $m_i(t, s)$ may no longer be an $\mathcal{F}_t(s)$ martingale since for the i-th subject with $s < U_i < t$, the enrollment allocation depends on the information up to its entry time U_i, that is $\mathcal{F}_{U_i-}$, which may not be contained in $\mathcal{F}_t(s)$. Similarly, the empirical process theory, which requires the independent allocation scheme, is also not applicable.

Examples of adaptive design/allocation schemes include the randomized play-the-winner rule (Wei and Durham, 1978), dynamic treatment regimes (Pocock and Simon, 1975, Robins, 1986, and Murphy and Bingham, 2008) with survival endpoints, efficient randomized adaptive designs (Hu, et al., 2009), and adaptive design with sample size re-estimation (Cui, et al., 1999 and Shen and Fisher, 1999).

3. Main Convergence Results.

For simplicity of notation, we assume $\mathcal{T} = \infty$ below. Following Sellke and Siegmund (1983) and Slud (1984), we introduce an index n to parameterize the size of the clinical trial. Thus, notation in Section 2 will include subscript n. Specifically, we have $R_{n,u}$ for R_u, $M_{n,t}$ for M_t, $p_n(.)$ for $p(.)$, $q_n(.)$ for $q(.)$, and $f_{n;s,u,z,\delta}(t)$ for $f_{s,u,z,\delta}(t)$. Additional quantities with the subscript n introduced henceforth are self-explained.

Let $[0, \mathcal{T}) \times \mathcal{X}_1$ with $\mathcal{X}_1 \subset \mathcal{D}_z \times \{0, 1\}$ be the sub-mark space in which we are interested. Note that for the Cox model, $\mathcal{X}_1 = \mathcal{D}_z \times \{1\}$ and $f_{n;s,u,z,\delta}(t) =$

$z(s-u) - \bar{Z}_n(\beta; t, s-u)$. Let

$$V_{n,t,\vartheta} \triangleq \int_0^t \int_0^{s\wedge\vartheta} \int_{\mathcal{X}_1} f_{n;s,u,z,\delta}^{\otimes 2}(t) q_n(ds\, du\, dz\, d\delta),$$

where $\vartheta \leq t$ are the entry time and the calendar time, respectively, and $a^{\otimes 2} = aa'$ for a column vector a. This may be interpreted as the accumulated information up to time t from all subjects whose entry times are before ϑ. Note that when $f_{n;s,u,z,\delta}(t)$ is $\mathcal{F}_{n,s}$ predictable, $V_{n,t,\vartheta}$ is the predictable variation process $\langle M^f([0,\vartheta] \times \mathcal{X}_1)\rangle(t)$ defined in (2.10). A natural estimator is

$$\hat{V}_{n,t,\vartheta} \triangleq \int_0^t \int_0^{s\wedge\vartheta} \int_{\mathcal{X}_1} f_{n;s,u,z,\delta}^{\otimes 2}(t) p_n(ds\, du\, dz\, d\delta).$$

For notational simplicity, when $\vartheta = t$, we will use $V_{n,t}$ and $\hat{V}_{n,t}$ for $V_{n,t,t}$ and $\hat{V}_{n,t,t}$ whenever there is no ambiguity.

3.1. *Uni-dimensional f_n.* In this subsection, we consider the case in which f_n is real-valued. It is well-known that in a clinical trial with survival as the endpoint, the power is associated with sample size through information accumulated within the study period; see Friedman, Furberg and DeMets (1998). Let V_n be the total information used in the study design. For the log-rank score process, Slud (1984) uses the number of enrollments for V_n, with n being the process index, while Sellke and Siegmund (1983) simply take $V_n = n$. In general, as t increases, the actual information $V_{n,t}$ increases and may reach a planned portion of the information V_n. However, due to interim adjustment which is common in adaptive designs, the ratio $V_{n,t}/V_n$ may not converge, making standard martingale central limit theorem not applicable for $M_t^f/\sqrt{V_n}$ in time scale t. To circumvent this difficulty, we will adopt an information-based time rescaling approach (Lai and Siegmund, 1983). Specifically, let

$$\sigma_{n,v} \triangleq \inf\{t : V_{n,t}/V_n \geq v\},$$

an estimator of which is

$$\hat{\sigma}_{n,v} \triangleq \inf\{t : \hat{V}_{n,t}/V_n \geq v\}.$$

We can interpret $\sigma_{n,v}$ as the calendar time at which a v proportion of the planned information has been accumulated. It is natural to expect that if $\sigma_{n,v} < \infty$, then

$$B_n(v) \triangleq \frac{1}{\sqrt{V_n}} \int_0^{\sigma_{n,v}} \int_0^s \int_{\mathcal{X}_1} f_{n;s,u,z,\delta}(\sigma_{n,v}) dM_{n,s}$$

will converge to the Brownian motion process. If $\hat{\sigma}_{n,v}$ is a consistent estimator, then we expect that

$$\hat{B}_n(v) \triangleq \frac{1}{\sqrt{V_n}} \int_0^{\hat{\sigma}_{n,v}} \int_0^s \int_{\mathcal{X}_1} f_{n;s,u,z,\delta}(\hat{\sigma}_{n,v}) dM_{n,s}$$

also converges to the Brownian motion. The following conditions are needed for the above stated Brownian approximation.

Condition B. For enrollment process R_n and information scale V_n, we require the following to be true:
(i) Total information $V_n \to \infty$ in probability as $n \to \infty$, and for any $t > 0$,

$$P(\tilde{T}_i \in [t, t+dt), \Delta_i | \mathcal{F}_{n,t-}, V_n) = P(\tilde{T}_i \in [t, t+dt), \Delta_i | \mathcal{F}_{n,t-}). \tag{3.1}$$

(ii) For any finite t and $\bar{v}$, $R_{n,t}/V_n < \infty$, $a.s.$ and $P(\sigma_{n,\bar{v}} < \infty) \xrightarrow{P} 1$ as $n \to \infty$.
(iii) For any $\tau < \infty$ and $0 < s, u < \tau$, there exists a constant $\tilde{K}_\tau$ such that for any Borel sets: $A \subset [0, \tau]$, $I \subset [0, \tau]$,

$$\lim_{n\to\infty} P\biggl(\sup_{A,I} \biggl\{ \int_A \int_I \int_{\mathcal{X}_1} q_n(ds\, du\, dz\, d\delta) - \tilde{K}_\tau \int_A \int_I ds\, dR_{n,u} \biggr\} < 0 \biggr) = 1.$$

Condition C. Score function $f_{n;s,u,z,\delta}(t)$ satisfies:
(i) $f_{n;s,u,z,\delta}(t) = f_{n;s',u',z',\delta}(t)$ if $s - u = s' - u'$ and $z(s-u) = z'(s'-u')$.
(ii) For any $\tau < \infty$, there exists a constant K_τ such that for $t \le \tau$

$$\lim_{n\to\infty} P\biggl(\sup_{\substack{0 \le t \le \tau \\ U_i \le t}} \biggl\{ |f_{n;U_i,U_i,Z_i,\Delta_i}(t)| + \int_0^{t-U_i} |f_{n;U_i+w,U_i,Z_i,\Delta_i}(t)| dw \biggr\} < K_\tau \biggr) = 1,$$

where $|\cdot|$ denotes the L_1-norm for a vector when f_n is multidimensional.
(iii) There exists an $\mathcal{F}_{n,s}$ predictable function $g_{n;s,u}(s)$ such that for $0 \le u, s \le t \le \tau$, $f_{n;s,u,Z_u,\Delta_u}(t) - g_{n;s,u}(s) \xrightarrow{P} 0$ uniformly as $n \to \infty$.

REMARK 3.1 (Condition B). *Equation (3.1) is trivially satisfied in commonly encountered cases since V_n is usually determined at the beginning of a trial. Part (ii) means that the sample size up to any time t cannot be more than a multiple of the planned information V_n, and that the time it takes to reach $\bar{v}V_n$ is finite. It is easy to see that in the case of the Cox model, this assumption is implied by Condition 4.1 of Sellke and Siegmund (1983). Part (iii) means intuitively that $q_n(ds\, du)/(ds\, dR_u) < \tilde{K}$ uniformly in probability. Again, for the Cox model, it is satisfied when the covariates and the baseline hazard functions are bounded.*

REMARK 3.2 (Condition C). *Because $s-u$ corresponds to survival time, part (i) is natural in that it requires the score and covariate functions to be on the survival time scale; see (2.14) as an example. Moreover, if (i) is not satisfied, we can construct a counterexample for Lemma 3.6 below, by letting $f_{n;s,u,z,\delta}(t) - g_{n;s,u}(s) = 1$ at the jump points of $M_{n,s}$ and 0 otherwise. Parts (ii) and (iii) are standard and analogous to Conditions 1-3 in Bilias et al. (1997).*

REMARK 3.3. *In practice, there might be a planned final analysis time τ_n. For instance, τ_n may be the time at which there are n events observed or at which a budget cap is reached. In this case, the stopping time $\sigma_{n,\bar{v}}$ can be still achieved before τ_n by taking a weight function w_n adaptively and by defining $w_n \cdot f_n$ as the new integrand function; see Shen and Cai (2003) for an example related to sample size reestimation. Therefore, Conditions B and C may still be satisfied.*

We now state the main result for uni-dimensional f_n. For any constant $\bar{v}$, let $D([0, \bar{v}])$ be the space of cadlag (right continuous with left limit) functions on $[0, \bar{v}]$ with the Skorokhod topology. Then, we have the following Brownian approximation of the score process. It extends the results of Sellke and Siegmund (1983) and Slud (1984) to cover the case with dependent entry times and a more general integrand function f_n.

THEOREM 3.4. *Under Conditions A, B, and C, we have the following weak convergence on the space $D([0, \bar{v}])$,*

$$\{B_n(v), 0 \le v \le \bar{v}\} \xrightarrow{\mathcal{D}} \{B(v), 0 \le v \le \bar{v}\},$$

where $B(v)$ is the Brownian motion process. Moreover, the convergence still holds with $\sigma_{n,v}$ replaced by $\hat{\sigma}_{n,v}$, i.e.,

$$\{\hat{B}_n(v), 0 \le v \le \bar{v}\} \xrightarrow{\mathcal{D}} \{B(v), 0 \le v \le \bar{v}\}.$$

Our proof of Theorem 3.4 relies on martingale-based techniques by making use of the martingale central limit theorem and certain maximal inequalities. It consists of several major steps corresponding to the following lemmas, whose proofs are given in Section 6.

When $f_{n;s,u,z,\delta}(t)$ is $\mathcal{F}_{n,s}$ predictable, results for martingales may be used and the martingale central limit theorem (Rebolledo, 1980) implies the following weak convergence.

LEMMA 3.5. *Suppose that $f_{n;s,u,z,\delta}(t)$ is $\mathcal{F}_{n,s}$ predictable and uniformly bounded in probability. Then under Condition B, we have*

$$\{B_n(v), 0 \le v \le \bar{v}\} \xrightarrow{\mathcal{D}} \{B(v), 0 \le v \le \bar{v}\}. \tag{3.2}$$

Moreover, the convergence continues to hold with $\sigma_{n,v}$ replaced by $\hat{\sigma}_{n,v}$, i.e.,

$$\{\hat{B}_n(v), 0 \le v \le \bar{v}\} \xrightarrow{\mathcal{D}} \{B(v), 0 \le v \le \bar{v}\}. \tag{3.3}$$

In general, with staggered entry, $f_{n;s,u,z,\delta}(t)$ is often not $\mathcal{F}_{n,s}$ predictable. Consequently, the corresponding integral may no longer be a martingale and Lemma 3.5 cannot be applied directly. The following lemma shows that it can be approximated by a martingale under suitable conditions.

LEMMA 3.6. *Let $\tau < \infty$. Under Conditions A, B, and C, we have*

$$\sup_{\vartheta,t\in[0,\tau]} \frac{1}{\sqrt{V_n}}\Big|\int_0^t \int_0^{s\wedge\vartheta}\int_{\mathcal{X}_1} [f_{n;s,u,z,\delta}(t) - g_{n;s,u}(s)]dM_{n,s}\Big| \xrightarrow{P} 0,$$

where $g_{n;s,u}(s)$ is defined as in Condition C(iii).

REMARK 3.7. *Lemma 3.6 provides a tightness result similar to that of Lemma 2 in Gu and Lai (1991). Our use of the martingale structure along the calendar time allows us to apply the martingale central theorem, bypassing empirical process based approximations that may not be applicable under adaptive design.*

PROOF OF THEOREM 3.4. With the preceding lemmas, it is now straightforward to prove Theorem 3.4. In view of Condition B, we only need to consider the case when $\sigma_{n,\bar{v}} < \tau$ a.s. with τ being a big enough constant. From Lemma 3.6, it suffices to show that for $0 \le v \le \bar{v}$,

$$\frac{1}{\sqrt{V_n}}\int_0^{\hat{\sigma}_{n,v}}\int_0^s\int_{\mathcal{X}_1} g_{n;s,u}(s)dM_{n,s}$$

converges weakly to the Brownian motion. From an argument similar to the proof of (3.3) (see Subsection 6.2.2), it is sufficient to show the weak convergence of

$$\frac{1}{\sqrt{V_n}}\int_0^{\hat{\sigma}'_{n,v}}\int_0^s\int_{\mathcal{X}_1} g_{n;s,u}(s)dM_{n,s},$$

where $\hat{\sigma}'_{n,v} = \inf\{t : \int_0^t\int_0^s\int_{\mathcal{X}_1} g^2_{n;s,u}(s)p_n(ds\,du\,dz\,d\delta) \ge vV_n\}$. Since $g_{n;s,u}(s)$ is $\mathcal{F}_{n,s}$ predictable, we get the desired conclusion from Lemmas 3.5. $\square$

3.2. *Multidimensional f_n.* For the multidimensional case, the above time rescaling approach may not be directly applicable since we cannot scale $V_{n,t,\vartheta}$ with a single growth rate in t. In other words, $\sigma_{n,t,\vartheta}$ is not well defined. Nevertheless, under the usual variance stability condition (see equation 3.4 below), we still have the weak convergence result, which extends Gu and Lai (1991) and Bilias, et al. (1997). More details in the case of the Cox model are given in Section 4.

THEOREM 3.8. *Let $\tau < \infty$ and assume there exists a nonrandom matrix function $V(t,\vartheta)$, such that for $0 < \vartheta \leq t \leq \tau$,*

$$\frac{V_{n,t,\vartheta}}{n} \xrightarrow{P} V(t,\vartheta). \tag{3.4}$$

Then under Conditions A, B(iii), and C, $n^{-1/2}\int_0^t \int_0^{s\wedge\vartheta} \int_{\mathcal{X}_1} f_{n;s,u,z,\delta}(t) dM_{n,s}$ converges to a zero-mean Gaussian process $\tilde{\xi}(t,\vartheta)$ on $\{t,\vartheta : 0 < \vartheta \leq t \leq \tau\}$ with continuous sample path and covariance function

$$E[\tilde{\xi}(t_1,\vartheta_1)\tilde{\xi}'(t_2,\vartheta_2)] = V(t_1 \wedge t_2, \vartheta_1 \wedge \vartheta_2).$$

PROOF OF THEOREM 3.8. In view of the proof of Theorem 3.4, a key step for obtaining the desired weak convergence result is to establish the tightness result analogous to Lemma 3.6. This is shown by the next lemma, whose proof is given in Subsection 6.3.

LEMMA 3.9. *Under the same assumptions as those of Theorem 3.8, we have*

$$\sup_{\vartheta,t\in[0,\tau]} \frac{1}{\sqrt{n}}\Big| \int_0^t \int_0^{s\wedge\vartheta} \int_{\mathcal{X}_1} f_{n;s,u,z,\delta}(t) - g_{n;s,u}(s) dM_{n,s}\Big| \xrightarrow{P} 0.$$

Then, Theorem 3.8 follows from Lemma 3.9 and the functional martingale central limit theorem. □

4. Cox Model with adaptive entry. In clinical trials with adaptive allocation rules, patients are accrued sequentially and treatment assignment may depend on the observed responses, leading to dependent enrollment processes. In this section, we use the marked point process framework to formulate the Cox proportional hazards model based score processes under response and/or covariate adaptive allocation schemes. We also discuss how the general results and conditions presented in Section 3 may be applied and verified under the Cox model.

4.1. *Cox model with unidimensional parameter.* Following the notation in Sections 2 and 3, we have the compensator $q_n(ds\,du\,dz\,d\delta)$ defined on $[0,\mathcal{T})\times\mathcal{X}$ satisfying

$$1(\delta=1)q_n(ds\,du\,dz\,d\delta) = \begin{cases} 1(\tilde{T}_u \geq s-u, Z_u \in dz)\exp\{\beta Z_u(s-u)\}\lambda_0(s-u)dR_{n,u}ds & s\geq u;\delta=1 \\ 0 & \text{otherwise.}\end{cases}$$

From (2.12), for each $k=0,1,2$, $\vartheta>0$, and $w>0$, we have

$$\begin{aligned}\Gamma_{n,k}(\beta;\vartheta,w) &= \sum_{U_i\leq\vartheta} Z_i^k(w)\exp(\beta Z_i(w))1(\tilde{T}_i\geq w) \\ &= \int_0^\vartheta 1(\tilde{T}_u\geq w)Z_u^k(w)\exp\{\beta Z_u(w)\}dR_{n,u}. \end{aligned}\tag{4.1}$$

The score processes as defined in Subsection 2.2 become

$$\begin{aligned} U_n(\beta;t) &= \int_0^t\int_0^s\int_{\mathcal{X}_1}[Z_u(s-u)-\bar{Z}_n(\beta;t,s-u)]dM_{n,s}, \\ U_n(\beta;t,\vartheta) &= \int_0^t\int_0^{s\wedge\vartheta}\int_{\mathcal{X}_1}[Z_u(s-u)-\bar{Z}_n(\beta;t,s-u)]dM_{n,s}, \end{aligned}$$

where $\mathcal{X}_1=\mathcal{D}_z\times\{1\}$ and

$$\bar{Z}_n(\beta;t,w) = \frac{\Gamma_{n,1}(\beta;t-w,w)}{\Gamma_{n,0}(\beta;t-w,w)}.$$

The following conditions for the Cox model imply Condition C in Section 3.

C1. For every i, $Z_i(\cdot)$ is bounded and of uniformly bounded variation in the sense that for any constant $\tau<\infty$, there exists a nonrandom constant K_τ such that for any subject i,

$$\lim_{n\to\infty} P\left(\sup_i\left\{|Z_i(0)|+\int_0^\tau |Z_i(ds)|\right\}\leq K_\tau\right)=1$$

C2. For any w and ϑ, there exists a $\mathcal{F}_{n,w-}$ measurable random variable (or constant) $\bar{E}_{n,k}(\vartheta,w)$ such that

$$\frac{1}{R_{n,\vartheta}}\sum_{U_i\leq\vartheta} Z_i^k(w)\exp(\beta_0 Z_i(w))E\big[1(\tilde{T}_i\geq w)|\mathcal{F}_{n,\mathcal{T},U_i-}\big]-\bar{E}_{n,k}(\vartheta,w)\xrightarrow{P}0.$$

REMARK 4.1. *Condition C2 states the stability condition. For the special case of independent enrollment allocation, R_t is non-informative for $\tilde{T}$ and Condition C2 holds naturally under the new filtration $\{\mathcal{F}'_{n,t}, 0 < t \le \mathcal{T}\}$ defined by $\mathcal{F}'_{n,t} = \mathcal{F}_{n,t} \cup \sigma\{R_{n,s}, \forall s > 0\}$.*

THEOREM 4.2. *Suppose that V_n is chosen to be n and that $f_{n;s,u,z,\delta}(t) = z(s-u) - \bar{Z}_n(\beta; t, s-u)$. Let $\hat{\sigma}_{n,v}$ be defined as in Subsection 3.1. Then under Conditions A, B, C1, and C2, for any $\bar{v}$, $n^{-1/2}U_n(\beta_0; \hat{\sigma}_{n,v})$ converges weakly to the Brownian motion process in $v \in [0, \bar{v}]$, where β_0 is the true parameter value.*

PROOF. By Condition B(ii), we only need to consider the case when $\sigma_{n,\bar{v}} < \tau$, where τ is a large constant. Since Conditions A, B, and C(i) in Sections 2 and 3 are satisfied, in order to apply Theorem 3.4, it suffices to show that Conditions C(ii) and C(iii) hold, i.e., $\bar{Z}_n(\beta_0; t, w)$ is of bounded variation and converges to an $\mathcal{F}_{n,w}$ predictable $Z_{n,p}(\beta_0; w)$ uniformly. This is equivalent to uniform convergence of $\bar{Z}_n(\beta_0; \vartheta + w, w)$ with respect to ϑ and w.

On $\{w, \vartheta : w > 0,\ \vartheta > 0,\ w + \vartheta \le \tau\}$, from (4.1), $\Gamma_{n,k}(\beta_0; \vartheta, w)$ can be expressed as an integral with respect to $p_n(ds\, du\, dz\, d\delta)$:

$$\Gamma_{n,k}(\beta_0; \vartheta, w) = \int_0^{\vartheta} \int_u^{\infty} \int_{\mathcal{X}} z^k(w) \exp(\beta_0 z(w)) I(s - u \ge w) p_n(du\, ds\, dz\, d\delta).$$

For k=0, 1 and 2, we know that for $0 \le \vartheta \le \tau - w$,

$$\begin{aligned} M_{n,k}(\vartheta) &= \frac{1}{n}\Big(\Gamma_{n,k}(\beta_0; \vartheta, w) \\ &\quad - \int_0^{\vartheta} E\big[1(\tilde{T}_u \ge w) Z_u^k(w) \exp(\beta_0 Z_u(w)) | \mathcal{F}_{n,\mathcal{T},u-}\big] dR_u\Big), \end{aligned}$$

as processes in ϑ are $\mathcal{F}_{n,\mathcal{T},\vartheta}$ martingales. A simple application of Lenglart's inequality (Lemma 6.3) gives that for any w,

$$\sup_{0 \le \vartheta \le \tau} M_{n,k}(\vartheta) \xrightarrow{P} 0, \quad k = 0, 1, 2. \tag{4.2}$$

Let $\epsilon > 0$ and define stopping time $a = \inf\{\vartheta : R_{n,\vartheta}/n > \epsilon\}$. Condition

C1 implies that for any η with $0 < \eta < \epsilon$,

$$\begin{aligned}(4.3)\quad & \sup_{\substack{a\le\vartheta_1,\vartheta_2\le\tau \\ |R_{n,\vartheta_1}-R_{n,\vartheta_2}|\le\eta n}} \Bigg| \frac{1}{R_{n,\vartheta_1}} \sum_{U_i\le\vartheta_1} E\big[1(\tilde{T}_i \ge w) Z_i^k(w) \exp(\beta_0 Z_i(w)) | \mathcal{F}_{n,\mathcal{T},U_i-}\big] \\ & \qquad - \frac{1}{R_{n,\vartheta_2}} \sum_{U_i\le\vartheta_2} E\big[1(\tilde{T}_i \ge w) Z_i^k(w) \exp(\beta_0 Z_i(w)) | \mathcal{F}_{n,\mathcal{T},U_i-}\big] \Bigg| \\ \le\ & 2K_\tau^k \exp(\beta_0 K_\tau)\eta/\epsilon,\end{aligned}$$

which gives the tightness result for

$$\sum_{U_i\le\vartheta} E[1(\tilde{T}_i \ge w) Z_i^k(w) \exp(\beta_0 Z_i(w)) | \mathcal{F}_{n,\mathcal{T},U_i-}]/R_{n,\vartheta}$$

on $\{\vartheta : a \le \vartheta \le \tau\}$. Therefore (4.3) together with Conditions C1 and C2, implies

$$\begin{aligned}(4.4)\quad & \sup_{a\le\vartheta\le\tau} \Bigg| \frac{1}{R_{n,\vartheta}} \sum_{U_i\le\vartheta} E\big[1(\tilde{T}_i \ge w) Z_i^k(w) \exp(\beta_0 Z_i(w)) | \mathcal{F}_{n,\mathcal{T},U_i-}\big] \\ & \qquad - \bar{E}_{n,k}(\vartheta, w) \Bigg| \xrightarrow{P} 0.\end{aligned}$$

From (4.2) and (4.4), for any w, as $n \to \infty$,

$$\sup_{a\le\vartheta\le\tau} \left| \frac{\Gamma_{n,k}(\beta_0;\vartheta,w)}{R_{n,\vartheta}} - \bar{E}_{n,k}(\vartheta,w) \right| \xrightarrow{P} 0.$$

Since $\exp(\beta Z_u(w))$ and $Z_u(w)\exp(\beta Z_u(w))$ are of bounded variation for $w \in [0,\tau]$, so is $\sup_{a\le\vartheta\le\tau} \Gamma_{n,k}(\beta_0;\vartheta,w)/R_{n,\vartheta}$. Therefore

$$\sup_{w} \sup_{a\le\vartheta\le\tau} \left| \frac{\Gamma_{n,k}(\beta_0;\vartheta,w)}{R_{n,\vartheta}} - \bar{E}_{n,k}(\vartheta,w) \right| \xrightarrow{P} 0,$$

which indicates the uniform convergence of $\bar{Z}_n(\beta_0;\vartheta + w, w)$ to an $\mathcal{F}_{n,w}$ predictable $Z_{n,p}(\beta_0;w)$ on $\{w,\vartheta : w + \vartheta \le \tau, \vartheta \ge a, w > 0\}$.

For $\vartheta > 0$, to show the bounded variation of $\bar{Z}_n$ in w, let $0 = \omega_1 < \cdots <$

$w_{n_0} = \tau$ be a partition of $[0, \tau]$. For $1 \le i < n_0$,

$$\begin{aligned}
&\left|\frac{\Gamma_{n,1}(\beta_0;\vartheta,w_{i+1})}{\Gamma_{n,0}(\beta_0;\vartheta,w_{i+1})} - \frac{\Gamma_{n,1}(\beta_0;\vartheta,w_i)}{\Gamma_{n,0}(\beta_0;\vartheta,w_i)}\right| \\
\le\ & \frac{|\Gamma_{n,1}(\beta_0;\vartheta,w_i) - \Gamma_{n,1}(\beta_0;\vartheta,w_{i+1})|}{\Gamma_{n,0}(\beta_0;\vartheta,w_i)} \\
&+\frac{|\Gamma_{n,0}(\beta_0;\vartheta,w_i) - \Gamma_{n,0}(\beta_0;\vartheta,w_{i+1})| \cdot |\Gamma_{n,1}(\beta_0;\vartheta,w_{i+1})|}{\Gamma_{n,0}(\beta_0;\vartheta,w_i)\Gamma_{n,0}(\beta_0;\vartheta,w_{i+1})} \\
\le\ & \frac{|\Gamma_{n,1}(\beta_0;\vartheta,w_i) - \Gamma_{n,1}(\beta_0;\vartheta,w_{i+1})|}{R_{n,\vartheta}} K'_\tau \\
&+\frac{|\Gamma_{n,0}(\beta_0;\vartheta,w_i) - \Gamma_{n,0}(\beta_0;\vartheta,w_{i+1})|}{R_{n,\vartheta}} K''_\tau,
\end{aligned}$$

where K'_τ and K''_τ are constant depending only on τ. The desired conclusion then follows from Condition C1.

Let $B_n(v) = n^{-1/2}U_n(\beta_0, \hat\sigma_{n,v})$. Since f_n^2 is bounded, say by constant M, we have $\int_0^a f_n^2 dp_n/n < M\epsilon$. Take ϵ small enough and let $\epsilon' = M\epsilon$; then $\hat\sigma_{n,\epsilon'} > a$. Therefore, in view of Lemmas 3.5 and 3.6, we get

$$\{B_n(v), \epsilon' \le v \le \bar v\} \xrightarrow{\mathcal{D}} \{B(v), \epsilon' \le v \le \bar v\}.$$

For the convergence property of the tail part $\{v : 0 < v \le \epsilon'\}$, we only need to show that for any $\eta,\ \eta_1 > 0$ and all big n,

$$P(\sup_{v \le \epsilon'} |B_n(v) - B(v)| > \eta_1) \le \eta.$$

Since $B(v)$ can be bounded near 0, it suffices to show

$$P(\sup_{v \le \epsilon'} |B_n(v)| > \eta_1/2) < \eta/2.$$

This can be done using an integration by parts argument similar to that of the proof of Lemma 3.6. $\square$

4.2. *Cox model with multidimensional regression parameter.* For a p-dimensional covariate vector Z and corresponding vector $\boldsymbol\beta$, the notation and conditions in Subsection 4.1 generalize naturally to the multidimensional case. Specifically, we use $Z^{\otimes k}$ for Z^k, where $Z^{\otimes 0} = 1$, $Z^{\otimes 1} = Z$, and $Z^{\otimes 2} = ZZ'$, and we take the $|\cdot|$ in Condition C1 as the L_1 norm for a p-dimensional vector. Under these modifications, $\Gamma_{n,k}(\boldsymbol\beta;\vartheta,w)$ becomes

$$\int_0^\vartheta 1(\tilde T_u \ge w) Z_u^{\otimes k}(w) \exp\{\boldsymbol\beta' Z_u(w)\} dR_{n,u}. \tag{4.5}$$

Additional quantities henceforth are self-explained.

To derive the weak convergence result analogous to Theorem 3.8, we need the following condition; see Conditions 2 and 3 in Bilias et al. (1997).

C3. For each $k = 0$, 1, and 2, there exists a non-random $\bar{E}_k(\vartheta, w)$ such that $\bar{E}_1(\cdot, w)/\bar{E}_0(\cdot, w)$ is continuous on $[0, \tau - w]$ and as $n \to \infty$,

$$\frac{R_{n,\vartheta}}{n} \bar{E}_{n,k}(\vartheta, w) - \bar{E}_k(\vartheta, w) \xrightarrow{P} 0,$$

for all positive ϑ, w satisfying $\vartheta + w \le \tau$, and

$$\sup_{0 \le t \le \tau} \int_0^t \left[\frac{\bar{E}_{n,1}(\vartheta, t - \vartheta)}{\bar{E}_{n,0}(\vartheta, t - \vartheta)} - \frac{\bar{E}_1(\vartheta, t - \vartheta)}{\bar{E}_0(\vartheta, t - \vartheta)} \right]^2 d\vartheta \to 0,$$

THEOREM 4.3. *Under Conditions A, C1 - C3, $n^{-1/2} U_n(\boldsymbol{\beta}_0; t)$ converges weakly to a vector-valued zero-mean Gaussian process ξ on $[0, \tau]$ with continuous sample path and covariance function $E[\xi(t_1)\xi'(t_2)]$ equal to*

$$\int_0^{t_1 \wedge t_2} \left[\bar{E}_2(t_1 \wedge t_2 - w, w) - \frac{\bar{E}_1^{\otimes 2}(t_1 \wedge t_2 - w, w)}{\bar{E}_0(t_1 \wedge t_2 - w, w)} \right] \lambda_0(w) dw.$$

Moreover, $n^{-1/2} U_n(\boldsymbol{\beta}_0; t, \vartheta)$ converges weakly to a vector-valued zero-mean Gaussian process $\tilde{\xi}(t, \vartheta)$ on $\{t, \vartheta : 0 \le \vartheta \le t \le \tau\}$ with continuous sample path and covariance function $E[\tilde{\xi}(t_1, u_1)\tilde{\xi}'(t_2, u_2)]$ equal to

$$\int_0^{t_1 \wedge t_2} \left[\bar{E}_2(u_{t_1 \wedge t_2, w}, w) - \frac{\bar{E}_1^{\otimes 2}(u_{t_1 \wedge t_2, w}, w)}{\bar{E}_0(u_{t_1 \wedge t_2, w}, w)} \right] \lambda_0(w) dw,$$

where $u_{t_1 \wedge t_2, w} = u_1 \wedge u_2 \wedge (t_1 \wedge t_2 - w)$.

PROOF. Thanks to Lemma 3.9, it suffices to prove that for every $n_0 > 0$ and partition $0 \le u_1 < \cdots < u_{n_0} \le \tau$, $\{n^{-1/2}\tilde{U}_n(\beta_0; t, u_1), \cdots, n^{-1/2}\tilde{U}_n(\beta_0; t, u_{n_0}), 0 \le t \le \tau\}$ converges weakly to a multivariate normal distribution $\{\tilde{\xi}(t, u_1), \cdots, \tilde{\xi}(t, u_{n_0})\}$, where

$$\tilde{U}_n(\boldsymbol{\beta}; t, \vartheta) = \int_0^t \int_0^{\vartheta \wedge s} \int_{\mathcal{X}_1} \left[Z_u(s - u) - \frac{\bar{E}_{n,1}(t - (s - u), s - u)}{\bar{E}_{n,0}(t - (s - u), s - u)} \right] M_n(ds).$$

From Lemma 2.2, $\{\tilde{U}_n(\boldsymbol{\beta}_0; t, u_j), \mathcal{F}_{n,t}, 0 \le t \le \tau\}$ are martingales with predictable covariation processes

$$\begin{aligned}
&\langle n^{-1/2}\tilde{U}_n(\boldsymbol{\beta}_0;\cdot,u_i), n^{-1/2}\tilde{U}_n(\boldsymbol{\beta}_0;\cdot,u_j)\rangle(t)\\
= \quad & \frac{1}{n}\int_0^t\int_0^{u_i\wedge u_j\wedge s}\int_{\mathcal{X}_1}\Big[Z_u(s-u)\\
&\qquad - \frac{\bar{E}_{n,1}(t-(s-u),s-u)}{\bar{E}_{n,0}(t-(s-u),s-u)}\Big]^{\otimes 2} q_n(ds\,du\,dz\,d\delta)\\
\xrightarrow{P} \quad & \int_0^t\Big[\bar{E}_2(u_{t,w},w)-\frac{\bar{E}_1^{\otimes 2}(u_{t,w},w)}{\bar{E}_0(u_{t,w},w)}\Big]\lambda_0(w)dw,
\end{aligned}$$

where $u_{t,w}=u_1\wedge u_2\wedge(t-w)$ and the convergence follows from Condition C3. Then, we can apply Rebolledo's (1980) martingale functional central limit theorem to obtain the desired weak convergence result in the same way as in the proof of Lemma 3.5. □

4.3. *Convergence of the maximum partial likelihood estimator.* Let the Cox partial likelihood estimator $\hat{\boldsymbol{\beta}}(t,\vartheta)$ and $\hat{\beta}(v)$ be solutions to $U(\boldsymbol{\beta};t,\vartheta)=0$ and $U(\beta;\hat{\sigma}_{n,v})=0$, respectively. In this subsection we give uniform consistency and weak convergence for the sequentially computed maximum partial likelihood estimator $\hat{\boldsymbol{\beta}}$ and $\hat{\beta}$.

We need the following condition, which ensures the presence of enough information, to gain the uniform consistency:

C4. There exists $\tau_0\in(0,\tau]$ such that

$$\liminf_{n\to\infty}\lambda_{min}\bigg(\frac{1}{n}\int_0^{\tau_0}\int_0^s\int_{\mathcal{X}_1}\Big[Z_u(s-u)-\frac{\bar{E}_{n,1}(\tau_0-(s-u),s-u)}{\bar{E}_{n,0}(\tau_0-(s-u),s-u)}\Big]^{\otimes 2} q_n(ds\,du\,dz\,d\delta)\bigg)=v_0>0,\ \ a.s.,$$

where $\bar{E}_{n,k}$ is defined as in Condition C2 and $\lambda_{min}(A)$ denotes the minimum eigenvalue of a symmetric matrix A.

Theorem 4.4 (One dimensional Cox model). *Under Condition C4 and the same assumptions as those of Theorem 4.2, for any $\bar{v}$, $\hat{\beta}(v)$ is uniformly consistent in the sense that*

$$\lim_{n\to\infty}\sup_{v_0\le v\le\bar{v}}|\hat{\beta}(v)-\beta_0|=0,\ \ in\ prob.$$

Moreover, $\{\sqrt{n}v(\hat{\beta}(v)-\beta_0), v_0\le v\le\bar{v}\}$ converges weakly to the Brownian motion process $B(v)$.

THEOREM 4.5 (Multidimensional Cox model). *Let $\tau < \infty$. Under Conditions A, C1, C2, and C4, $\hat{\boldsymbol{\beta}}(t,\vartheta)$ is uniformly consistent in the sense that*

$$\lim_{n\to\infty} \sup_{\tau_0\le\vartheta\le t\le\tau} ||\hat{\boldsymbol{\beta}}(t,\vartheta) - \boldsymbol{\beta}_0|| = 0, \ \textit{in prob.}$$

Moreover, if C3 is also satisfied, $\{\sqrt{n}(\hat{\boldsymbol{\beta}}(t,\vartheta)-\boldsymbol{\beta}_0), 0 \le \vartheta \le t \le \tau\}$ converges weakly to a vector-valued zero-mean Gaussian random field η with covariance

$$E[\eta(t_1,u_1)\eta'(t_2,u_2)] = E^{-1}[\tilde{\xi}^{\otimes 2}(t_1,u_1)]E[\tilde{\xi}(t_1,u_1)\tilde{\xi}'(t_2,u_2)]E^{-1}[\tilde{\xi}^{\otimes 2}(t_2,u_2)],$$

where $\tilde{\xi}$ is defined as in Theorem 4.3.

To prove Theorem 4.5, we need the following Lemma, which is a restatement of Lemma A.5 in Bilias, et al. (1997).

LEMMA 4.6. *Consider a set of functions $\{f_{n,\alpha} : n \ge 1, \alpha \in A\}$ from R^d to R^d. Suppose that (i) $\frac{\partial}{\partial\theta} f_{n,\alpha}(\theta)$ are nonnegative definite for all n, α, θ; (ii) $\sup_\alpha ||f_{n,\alpha}(\theta_0)|| \to 0$ as $n \to \infty$; (iii) there exists a neighborhood of θ_0, denoted by $\mathcal{N}(\theta_0)$, such that*

$$\liminf_{n\to\infty} \inf_{\theta\in\mathcal{N}(\theta_0)} \inf_{\alpha\in A} \lambda_{min}\left(\frac{\partial f_{n,\alpha}(\theta)}{\partial\theta}\right) > 0,$$

where λ_{min} is the minimum eigenvalue as defined in C4. Then there exists n_0 such that for every $n > n_0$ and $\alpha \in A$, $f_{n,\alpha}$ has a unique root $\theta_{n,\alpha}$ and $\sup_{\alpha\in A} ||\theta_{n,\alpha} - \theta_0|| \to 0$.

PROOF OF THEOREM 4.5. We only need to prove the multidimensional case. From the same argument as in the proof of Theorem 4.2, as $n \to \infty$,

$$\sup_{\vartheta,t} \left|\frac{1}{n} U(\boldsymbol{\beta}_0; t, \vartheta)\right| \xrightarrow{P} 0,$$

and

$$\begin{aligned}(4.6)\qquad \sup_{\vartheta,t} \frac{1}{n}\bigg|\frac{\partial}{\partial\boldsymbol{\beta}} U(\boldsymbol{\beta}_0;t,\vartheta) - \int_0^t \int_0^{s\wedge\vartheta} \int_{\mathcal{X}_1} \Big[Z_u(s-u) \\ - \frac{\bar{E}_{n,1}(t-(s-u), s-u)}{\bar{E}_{n,0}(t-(s-u), s-u)}\Big]^{\otimes 2} q_n(ds\,du\,dz\,d\delta)\bigg| \xrightarrow{P} 0.\end{aligned}$$

Since $\frac{1}{n}\frac{\partial}{\partial\boldsymbol{\beta}} U(\boldsymbol{\beta}_0;t,\vartheta)$ has a uniformly bounded derivative with respect to $\boldsymbol{\beta}$, Condition C4 and (4.6) imply that there exists a neighborhood of $\boldsymbol{\beta}_0$, $\mathcal{N}(\boldsymbol{\beta}_0)$, such that in probability

$$(4.7)\quad \liminf_{n\to\infty} \inf_{\tau_0\le\vartheta\le t\le\tau} \inf_{\boldsymbol{\beta}\in\mathcal{N}(\boldsymbol{\beta}_0)} \lambda_{min}\Big(\frac{1}{n}\frac{\partial}{\partial\boldsymbol{\beta}} U(\boldsymbol{\beta}_0;t,\vartheta)\Big) \ge \frac{v_0}{2} > 0.$$

Therefore, by Lemma 4.6, we get consistence.

By the Taylor expansion, uniformly on $0 \le \vartheta \le t \le \tau$,

$$\begin{aligned} & \frac{1}{\sqrt{n}} U(\hat{\boldsymbol{\beta}}(t,\vartheta); t, \vartheta) \\ = \quad & \frac{1}{\sqrt{n}} U(\boldsymbol{\beta}_0; t, \vartheta) + \frac{1}{n} \frac{\partial}{\partial \boldsymbol{\beta}} U(\boldsymbol{\beta}_0; t, \vartheta) \sqrt{n} (\hat{\boldsymbol{\beta}}(t,\vartheta) - \boldsymbol{\beta}_0) + o_p(1), \end{aligned}$$

where o_p is uniform for $\tau_0 \le \vartheta \le t \le \tau$. From this and Theorem 4.3, the weak convergence of $\sqrt{n}(\hat{\boldsymbol{\beta}}(t,\vartheta) - \boldsymbol{\beta}_0)$ follows. □

5. Discussion. General statistical methods and theory usually assume observations from different study subjects are independent. In practice, such an assumption may be violated. This paper deals with survival studies in which patients' entry and treatment allocations are adaptive and dependent on previous outcomes. Through carefully defined marked point processes, it provides a general framework under which a martingale-based approach is developed. It is shown that the usual score process for sequential data monitoring (Jennison and Turnbull, 2000 and Proschan, Lan and Wittes, 2006) can still be approximated by a time rescaled Brownian motion process that is the theoretical cornerstone for modern group sequential methods for clinical trials. The results establish a bridge between sequential analysis for survival endpoints with staggered entry (Sellke and Siegmund, 1983, Slud, 1984, Gu and Lai 1991 and Bilias, et al. 1997) and covariate/response adaptive treatment allocation designs (Hu and Rosenberger, 2006). Specific details are given for the Cox model based score processes.

The theoretical framework and asymptotical results developed in this paper may be extended to other follow-up studies with more general outcome variables. For studies with longitudinal outcomes, dynamic regression models have been proposed and studied (see Martinussen and Scheike, 2000). Consideration of staggered entry and outcomes dependent allocation could complicate the analysis considerably. It is hoped that the approach of this paper will provide a basis for developing a new way to handle such study designs.

6. Proofs of the Main Results.

6.1. *Proof of Lemma 2.2.* Following Chapter II in Jacod and Shiryaev (2003), random measure $p(\cdot)$ has a $\mathcal{F}_t$ predictable compensator $q(\cdot)$ of form

$E(p(dt\,du\,dz\,d\delta)|\mathcal{F}_{t-})$, and

$$M_t^f(\mathcal{X}) = \int_0^t \int_{\mathcal{X}} f_{s,u,z,\delta}(s)(p(ds\,du\,dz\,d\delta) - q(ds\,du\,dz\,d\delta)) \tag{6.1}$$

is a martingale for any $\mathcal{F}_s$ predictable and integrable function $f_{s,u,z,\delta}(s)$. Thus, the first part of Lemma 1 follows.

As to the second part, for fixed t and any $0 \leq u_1 < u_2 \leq t$,

$$\begin{aligned} & E(M_{t,u_2}(E) - M_{t,u_1}(E)|\mathcal{F}_{t,u_1}) \\ = \;& E(\int_0^t \int_{(u_1,u_2]} \int_E (p-q)(ds\,du\,dz\,d\delta)|\mathcal{F}_{t,u_1}) \\ = \;& 0, \end{aligned}$$

which is due to Condition A and the first conclusion in this Lemma. Therefore the desired result follows.

6.2. *Proof of Lemma 3.5.* We separate our proof into two parts. In Subsection 6.2.1, we prove (3.2). (3.3) is proved in Subsection 6.2.2.

6.2.1. *Proof of* (3.2). Since $f_{n;s,u,z,\delta}(t)$ is assumed to be $\mathcal{F}_{n,s}$ predictable, denote $f_{n;s,u,z,\delta}(t)$ by $f_{n;s,u,z,\delta}(s)$ when there is no ambiguity. Consider the new filtration

$$\mathcal{F}'_{n,t} = \sigma\{\mathcal{F}_{n,t}, V_n\}.$$

Under this filtration, since R_u, Z_u are predictable,

$$\begin{aligned} & E(p_n(ds\,du\,dz\,d\delta)|\mathcal{F}'_{n,t}) \\ = \;& E(1(u + \tilde{T}_u \in ds, Z_u \in dz, \Delta_u = \delta)dR_u|\mathcal{F}'_{n,t}) \\ = \;& E(1(u + \tilde{T}_u \in ds, \Delta_u = \delta)|\mathcal{F}'_{n,t})1(Z_u \in dz)dR_u, \end{aligned}$$

which equals, by Condition B(i), the compensator for $p_n(ds\,du\,dz\,d\delta)$ under $\mathcal{F}_{n,t}$, i.e.,

$$q_n(ds\,du\,dz\,d\delta) = E(1(u + \tilde{T}_u \in ds, \Delta_u = \delta)|\mathcal{F}_{n,t})1(Z_u \in dz)dR_u.$$

Then $B_n(v)$ is a martingale with respect to our new filtration by Lemma 2.2.

Note that $\langle B_n \rangle(v) = v$. From the central limit theorem for martingales in Rebolledo (1980), we only need to show that the quadratic variation process of any ϵ jump process converges to zero in probability. For any $\epsilon > 0$, the ϵ jump process

$$\begin{aligned} & B_{n,\epsilon}(v) \\ = & \sum_{w \le v} \Delta B_n(w) 1\left(|\Delta B_n(w)| \ge \epsilon\right) \\ = & \frac{1}{\sqrt{V_n}} \int_0^{\sigma_v} \int_0^s \int_{\mathcal{X}_1} f_{n;s,u,z,\delta}(s) 1\left(\left| \frac{f_{n;s,u,z,\delta}(s)}{\sqrt{V_n}} \right| \ge \epsilon \right) p_n(ds\, du\, dz\, d\delta) \\ = & \frac{1}{\sqrt{V_n}} \int_0^{\sigma_v} \int_0^s \int_{\mathcal{X}_1} f_{n;s,u,z,\delta}(s) 1\left(\left| \frac{f_{n;s,u,z,\delta}(s)}{\sqrt{V_n}} \right| \ge \epsilon \right) dM_{n,s} \\ & + \frac{1}{\sqrt{V_n}} \int_0^{\sigma_v} \int_0^s \int_{\mathcal{X}_1} f_{n;s,u,z,\delta}(s) 1\left(\left| \frac{f_{n;s,u,z,\delta}(s)}{\sqrt{V_n}} \right| \ge \epsilon \right) q_n(ds\, du\, dz\, d\delta). \end{aligned}$$

Thus, as f_n is uniformly bounded in probability

$$\begin{aligned} & \langle B_{n,\epsilon} \rangle(v) \\ = & \frac{1}{V_n} \int_0^{\sigma_v} \int_0^s \int_{\mathcal{X}_1} f^2_{n;s,u,z,\delta}(s) 1\left(\left| \frac{f_{n;s,u,z,\delta}(s)}{\sqrt{V_n}} \right| \ge \epsilon \right) q_n(ds\, du\, dz\, d\delta) \\ \xrightarrow{P} & 0, \quad as\ n \to \infty. \end{aligned}$$

Therefore $\{B_n(v), 0 \le v \le \bar{v}\} \xrightarrow{\mathcal{D}} \{B(v), 0 \le v \le \bar{v}\}$.

6.2.2. *Proof of* (3.3). Similarly we denote $f_{n;s,u,z,\delta}(t)$ by $f_{n;s,u,z,\delta}(s)$, which is $\mathcal{F}_{n,s}$ predictable. By Condition B (ii), we only need to consider the case when $\sigma_{n,\bar{v}}, \hat{\sigma}_{n,\bar{v}} < \tau$ for some constant $\tau < \infty$.

Martingale

$$m_n(t) \triangleq \frac{1}{V_n} \int_0^t \int_{\mathcal{X}} f^2_{n;s,u,z,\delta}(s) \big(p_n(ds\, du\, dz\, d\delta) - q_n(ds\, du\, dz\, d\delta) \big) \tag{6.2}$$

satisfies $\langle m_n(\tau) \rangle \xrightarrow{P} 0$ for uniformly bounded f_n under Condition B. From Lenglart's inequality, we get

$$\sup_{0 \le t \le \tau} m_n(t) \xrightarrow{P} 0.$$

Therefore for any $\delta > 0$, the definition of $\hat{\sigma}_{n,v}$ implies that the following holds uniformly in probability as $n \to \infty$

$$(v - \delta) V_n < \int_0^{\hat{\sigma}_{n,v}} \int_{\mathcal{X}} f^2_{n;s,u,z,\delta}(s) p_n(ds\, du\, dz\, d\delta) - m_n(\hat{\sigma}_{n,v}) \cdot V_n,$$

$$(v+\delta)V_n > \int_0^{\hat{\sigma}_{n,v}} \int_{\mathcal{X}} f^2_{n;s,u,z,\delta}(s) p_n(ds\,du\,dz\,d\delta) - m_n(\hat{\sigma}_{n,v}) \cdot V_n,$$

which indicates $\left| vV_n - \int_0^{\hat{\sigma}_{n,v}} \int_{\mathcal{X}} f^2_{n;s,u,z,\delta}(s) q_n(ds\,du\,dz\,d\delta)\right| < \delta V_n$. Therefore,

$$P(\sigma_{n,v-\delta} < \hat{\sigma}_{n,v} < \sigma_{n,v+\delta}, \forall v \in [0,\bar{v}]) \to 0 \text{ as } n \to \infty,$$

and the following holds uniformly in probability as $n \to \infty$

$$\sup_{0<v<\bar{v}} \left| B_n(v) - \hat{B}_n(v) \right| \le \sup_{\substack{0<s,t<\bar{v} \\ |s-t|<\delta}} \left| B_n(t) - B_n(s) \right|.$$

Therefore, from Lemma 3.5, for any ϵ, $\eta > \eta_1 > 0$, there exists $\delta > 0$ such that for big enough n,

$$P\left(\sup_{0<v<\bar{v}} |B_n(v) - \hat{B}_n(v)| > \epsilon \right) \le P\left(\sup_{\substack{0<s,t<\bar{v} \\ |s-t|<\delta}} |B_n(t) - B_n(s)| > \epsilon \right) + \eta_1 < \eta,$$

which completes our proof.

6.3. *Proof of Lemma 3.6 and Lemma 3.9.* We only need to prove Lemma 3.6; Lemma 3.9 follows from the same arguments. Since Δ is discrete, without loss of generality, it suffices to consider the sub-mark space $\mathcal{X}_1$ with $\delta = 1$. For statement convenience, we use $f_n(t,s,u)$ for $f_{n;s,u,Z_u,\Delta_u}(t)$ and $g_n(s,u)$ for $g_{n;s,u}(s)$.

Note that $p_n(ds\,du\,dz, \delta = 1) = 1(u + \tilde{T}_u \in ds, Z_u \in dz, \Delta_u = 1) dR_{n,u}$. From Condition A and Lemma 2.2,

$$\begin{aligned} q_n(ds\,du\,dz, \delta = 1) &= E(p_n(ds\,du\,dz, \delta = 1) | \mathcal{F}_{n,s-}) \\ &= E(1(u + \tilde{T}_u \in ds, \Delta_u = 1) | \mathcal{F}_{n,s-}) 1(Z_u \in dz) dR_{n,u}. \end{aligned}$$

From the above, we can define a new counting measure and its compensator on $[0,\tau] \times [0,\tau]$ by $p^*_n(ds\,du) = 1(u + \tilde{T}_u \in ds, \Delta_u = 1) dR_{n,u}$ and $q^*_n(ds\,du) = E(1(u + \tilde{T}_u \in ds, \Delta_u = 1) | \mathcal{F}_{n,s-}) dR_{n,u}$. Note that $p^*_n(ds\,du) = \int_{\mathcal{X}_1} p_n(ds\,du\,dz\,d\delta)$ and $q^*_n(ds\,du) = \int_{\mathcal{X}_1} q_n(ds\,du\,dz\,d\delta)$. From Lemma 2.2, we get the martingale measure $dM^*_{n,s} = p^*_n(ds\,du) - q^*_n(ds\,du)$, and for any $\mathcal{F}_{n,t}$ measurable and integrable f_n,

$$\int_0^t \int_0^{\vartheta} \int_{\mathcal{X}_1} f_{n;s,u,z,\delta}(t) dM_{n,s} = \int_0^t \int_0^{\vartheta} f_n(t,s,u) dM^*_{n,s}.$$

When there is no ambiguity, we use the notation, $p_n(ds\,du), q_n(ds\,du)$, and $M_{n,s}$, instead of $p^*_n(ds\,du)$, $q^*_n(ds\,du)$, and $M^*_{n,s}$.

Let $p_n(ds,u) = I(u + \tilde{T}_u \in ds)$ and $q_n(ds,u) = E(p_n(ds,u)|\mathcal{F}_{n,s-})$, which are the counting measure and the corresponding compensator for the subject who enrolled at time u. Then Lemma 2.2 implies that

$$M_n(ds,u) = p_n(ds,u) - q_n(ds,u)$$

is a martingale measure on $[u,\tau]$, which defines a basic martingale measure for each subject in the sense that if $u = U_i$, $M_n(ds,u) = I(U_i + \tilde{T}_i \in ds) - q_n(ds, U_i)$; see (2.16) as an example for the Cox model. Let $M_{n,t}(u) = \int_u^t M_n(ds,u)$, which is the total measure of interval $[u,t]$ under $M_n(ds,u)$. Let $M_{n,t}(du) = [\int_u^t M_n(ds,u)] \cdot dR_{n,u}$, which defines a martingale measure along entry time for all subjects who enrolled before time t.

Denote $M_{n,t,\vartheta}(\mathcal{X}_1)$ in Lemma 2.2 by $M_{n,t,\vartheta}$. It can be taken as a martingale along both calendar and entry times, i.e., $M_{n,t,\vartheta} = \int_0^t \int_0^{\vartheta\wedge s} dM_{n,s}$ is a martingale in t for any ϑ and $M_{n,t,\vartheta} = \int_0^{\vartheta} M_{n,t}(du)$ is a martingale in ϑ for any t. When $\vartheta = t$, we have $M_{n,t,t} = \int_0^t \int_0^s dM_{n,s}$, which is $M_{n,t}(\mathcal{X}_1)$. Similarly, define random integral $\tilde{M}_{n,w,\vartheta}$ with respect to survival time w and entry time ϑ by

$$\tilde{M}_{n,w,\vartheta} = \int_0^{\vartheta} M_{n,w+u}(du) \left(= \int_0^{\vartheta} M_{n,w+u}(u) dR_{n,u} \right).$$

Note that $\tilde{M}_{n,w,\vartheta}$ is defined on the information observed before entry time ϑ and survival time w.

We now proceed to prove Lemma 3.6. The following two propositions play a key role, and we will give their proofs in Sections 6.3.1 and 6.3.2, respectively. Proposition 6.1 shows the tightness for $M_{n,t,\vartheta}/\sqrt{V_n}$ along calendar and entry time.

PROPOSITION 6.1. *Under Conditions A and B, for any $\epsilon > 0$, there exist a constant $n_0 < \infty$ and partitions $0 = u_{n,0} \le u_{n,1} \le \cdots \le u_{n,n_0} = \tau$, which may be random, such that for all large n,*

$$P\left(\max_{0\le j<n_0} \sup_{\substack{\vartheta\in[u_{n,j},u_{n,j+1}];\\ 0\le t\le\tau}} |W_{n,t,\vartheta} - W_{n,t,u_{n,j}}| \ge \epsilon \right) \le \epsilon,$$

where $W_{n,t,\vartheta} = M_{n,t,\vartheta}/\sqrt{V_n}$.

The following proposition shows the tightness property for $\tilde{M}_{n,w,\vartheta}/\sqrt{V_n}$ along survival and entry time.

PROPOSITION 6.2. *Under Conditions A and B, for any $\epsilon > 0$, there exist partitions $0 = w_0 < w_1 < \cdots < w_{N_0} = \tau$ and $0 = u_{n,0} \le u_{n,1} \le \cdots \le u_{n,n_0} = \tau$ such that for all large n,*

$$P\biggl(\max_{\substack{0\le j<n_0\\0\le k<N_0}}\sup_{\substack{\vartheta\in[u_{n,j},u_{n,j+1}]\\w\in[w_k,w_{k+1}]}}|\tilde{W}_{n,w,\vartheta}-\tilde{W}_{n,w_k,u_{n,j}}|\ge\epsilon\biggr)\le\epsilon,$$

where $\tilde{W}_{n,w,\vartheta} = \tilde{M}_{n,w,\vartheta}/\sqrt{V_n}$.

PROOF OF LEMMA 3.6. To show the uniform convergence result, note that for any $0 \le \vartheta \le t \le \tau$,

$$\begin{aligned}
&\frac{1}{\sqrt{V_n}}\int_0^t\int_0^{s\wedge\vartheta}(f_n(t,s,u)-g_n(s,u))(p_n(ds\,du)-q_n(ds\,du))\\
=\;&\frac{1}{\sqrt{V_n}}\int_0^\vartheta\int_u^t(f_n(t,s,u)-g_n(s,u))(p_n(du\,ds)-q_n(du\,ds))\\
=\;&\frac{1}{\sqrt{V_n}}\int_0^\vartheta\Bigl[\int_u^t(f_n(t,s,u)-g_n(s,u))M_n(ds,u)\Bigr]dR_{n,u}.
\end{aligned}$$

Since the total variations of f_n and g_n are bounded and $M_{n,s,u}$ is a martingale with jump no bigger than 1, quadratic covariation $\bigl[f_n(t,\cdot,u)-g_n(\cdot,u),V_n^{-1/2}\cdot M_{n,\cdot,u}\bigr](t)$ converges to 0 uniformly in probability; then, using integration by parts, we have

$$\begin{aligned}
&\frac{1}{\sqrt{V_n}}\int_0^\vartheta\Bigl[\int_u^t(f_n(t,s,u)-g_n(s,u))M_n(ds,u)\Bigr]dR_{n,u}\\
=\;&\frac{1}{\sqrt{V_n}}\int_0^\vartheta\Bigl[M_{n,t}(u)(f_n(t,t,u)-g_n(t,u))-M_{n,u}(u)(f_n(t,u,u)-g_n(u,u))\\
&-\int_u^t M_{n,s}(u)(f_n(t,ds,u)-g_n(ds,u))\Bigr]dR_{n,u}+o_n(1)\\
=\;&\frac{1}{\sqrt{V_n}}\int_0^\vartheta(f_n(t,t,u)-g_n(t,u))M_{n,t}(du)\\
&-\frac{1}{\sqrt{V_n}}\int_0^\vartheta\Bigl[\int_u^t M_{n,s}(u)(f_n(t,ds,u)-g_n(ds,u))\Bigr]dR_{n,u}+o_n(1)
\end{aligned}$$

, (6.3)

where $o_n(1)$ is a small term converging uniformly to 0 in probability.

Consider the first term in (6.3). For any $\epsilon > 0$, Proposition 6.1 shows that there exists a partition $0 = u_0 \le u_{n,1} \le \cdots \le u_{n,n_0} = \tau$ such that for all large n, with probability bigger than $1-\epsilon$,

$$\sup_{i;u\in(u_{n,i},u_{n,i+1}]} |M_{n,t,u_{n,i+1}} - M_{n,t,u}|/\sqrt{V_n} < \epsilon.$$

Therefore, again by integration by parts, the following result holds uniformly on $0 \le \vartheta \le t \le \tau$ for all large n, with probability bigger than $1-2\epsilon$:

$$\begin{aligned}
& \left|\frac{1}{\sqrt{V_n}}\int_0^\vartheta (f_n(t,t,u)-g_n(t,u))M_{n,t}(du)\right| \\
= & \left|\frac{1}{\sqrt{V_n}}(f_n(t,t,\vartheta)-g_n(t,\vartheta))M_{n,t,\vartheta} - \frac{1}{\sqrt{V_n}}(f_n(t,t,0)-g_n(t,0))M_{n,t,0}\right. \\
& \left. -\frac{1}{\sqrt{V_n}}\int_0^\vartheta M_{n,t,u}(f_n(t,t,du)-g_n(t,du))\right| + o_n(1) \\
\le & \left|\frac{1}{\sqrt{V_n}}(f_n(t,t,\vartheta)-g_n(t,\vartheta))M_{n,t,\vartheta}\right| \\
& +\frac{1}{\sqrt{V_n}}\sum_{i=1}^{n_0}\left|\int_{u_{n,i-1}}^{u_{n,i}} M_{n,t,u_{n,i+1}}(f_n(t,t,du)-g_n(t,du))\right| + 2\epsilon\cdot K_\tau \\
\le & \ 3\epsilon\cdot K_\tau,
\end{aligned}$$

where K_τ is the total variation bound for $f_n(t,s,u) = f_n(t,s-u,0)$, and the last step follows from the Lenglart inequality.

For the second term in (6.3), by Condition C(i),

$$\begin{aligned}
& \frac{1}{\sqrt{V_n}}\int_0^\vartheta\int_u^t M_{n,s}(u)(f_n(t,ds,u)-g_n(ds,u))dR_{n,u} \\
= & \frac{1}{\sqrt{V_n}}\int_0^\vartheta\int_u^t M_{n,s}(u)(f_n(t,d(s-u),0)-g_n(d(s-u),0))dR_{n,u} \\
= & \frac{1}{\sqrt{V_n}}\int_0^t\Big[\int_0^{(t-w)\wedge\vartheta} M_{n,w+u}(u)dR_{n,u}\Big](f_n(t,dw,0)-g_n(dw,0)).
\end{aligned}$$

Recall that we let $\tilde{M}_{n,w,\vartheta} = \int_0^\vartheta M_{n,w+u}(u)dR_{n,u}$; then, from Proposition 6.2, there exist partitions $0 = w_0 < w_1 < \cdots < w_{N_0} = \tau$ and $0 = u_{n,0} \le u_{n,1} \le \cdots \le u_{n,n_0} = \tau$ such that

$$\sup_{\substack{i,j;w\in[w_i,w_{i+1}),\\ u\in[u_{n,j},u_{n,j+1})}} \frac{1}{\sqrt{V_n}}|\tilde{M}_{n,w,u} - \tilde{M}_{n,w_i,u_{n,j}}| < \epsilon;$$

then, similar to the proof for the first term in (6.3), we get that the following holds with probability bigger than $1-2\epsilon$ for all large n:

$$\begin{aligned}
&\frac{1}{\sqrt{V_n}}\int_0^t\Big[\int_0^{(t-w)\wedge\vartheta} M_{n,w+u}(u)dR_{n,u}\Big](f_n(t,dw,0)-g_n(dw,0))\\
\le\;& \frac{1}{\sqrt{V_n}}\sum_{i=1}^{N_0}\sum_{j=1}^{n_0}\left|\tilde{M}_{n,w_i,u_{n,j}}\int_{w_{i-1}}^{w_i}(f_n(t,dw,0)-g_n(dw,0))\right| + 2\epsilon\cdot K_\tau\\
\le\;& 3\epsilon K_\tau.
\end{aligned}$$

Therefore, combining the above inequalities, we have that for all large n,

$$P\left(\sup_{\vartheta,t\in[0,\tau]}\frac{1}{\sqrt{V_n}}\left|\int_0^t\int_0^{s\wedge\vartheta} f_n(t,s,u)-g_n(s,u)dM_{n,s}\right| < 6\epsilon K_\tau+\epsilon\right) > 1-5\epsilon,$$

which completes our proof. □

6.3.1. *Proof of Proposition 6.1.* We shall make use of some of the basic martingale inequalities given in the following lemma, which is due to Lenglart, Lepingle and Pratelli (1980).

Lemma 6.3. *Let $\{W(s),\mathcal{G}(s),s\ge 0\}$ be a martingale with right continuous paths and left limits. For any $q>1$, there exists a constant C_q depending only on q, such that*

$$E\left(\sup_{s\le\tau}|W(s)|^q\right)\le C_q\left(E[\langle W\rangle(\tau)]^{q/2}+E(\sup_{s\le\tau}|\triangle W(s)|^q)\right). \tag{6.4}$$

Moreover, if $\sup_{s\le\tau}|\triangle W(s)|\le c$, then for any $a,b>0$

$$P\left(\sup_{s\le\tau}|W(s)|\ge a,\langle W\rangle(\tau)\le b\right)\le 2\exp\left(-\frac{a^2}{2b}\psi(ac/b)\right),$$

where $\psi(x)=2x^{-2}\{(1+x)[\log(1+x)-1]+1\}$.

Proof of Proposition 6.1. Choose positive numbers p, $q>1$ such that $pq/2-p-q>1$. Let $u_0=0$ and define $u_{n,j}$ inductively by

$$u_{n,j+1}=\inf\{\vartheta:\vartheta>u_{n,j},2\tau\tilde{K}_\tau(R_{n,\vartheta}-R_{n,u_{n,j}})\ge\epsilon^pV_n\}\wedge(u_{n,j}+\epsilon^p)\wedge\tau.$$

Condition B(ii) implies that there are maximally $O(\epsilon^{-p})$ many, say n_0, distinct points in $[0,\tau]$ for all big n. From Lemma 1, $\{W_{n,t,\vartheta},\ \mathcal{F}_{n,t},\ t\ge 0\}$ is a martingale, and we know that $u_{n,j}$, $j = 1, \cdots, n_0$, are $\{\mathcal{F}_{n,t}, 0 < t \le \tau\}$ predictable. Thus, $\{\sup_{\vartheta\in[u_{n,j},u_{n,j+1}]} |W_{n,t,\vartheta} - W_{n,t,u_{n,j}}|, \mathcal{F}_{n,t}, t \ge 0\}$ is a nonnegative submartingale. By the Morkov inequality and Doob's (1953) maximal inequality,

$$
\begin{aligned}
& P\Bigg(\max_{0\le j<n_0} \sup_{\substack{\vartheta\in[u_{n,j},u_{n,j+1}];\\ 0\le t\le\tau}} |W_{n,t,\vartheta} - W_{n,t,u_{n,j}}| \ge \epsilon \Bigg) \\
\le\ & \frac{1}{\epsilon^q} \sum_{j=0}^{n_0-1} E\Bigg(\sup_{\substack{\vartheta\in[u_{n,j},u_{n,j+1}];\\ 0\le t\le\tau}} |W_{n,t,\vartheta} - W_{n,t,u_{n,j}}|^q \Bigg) \\
\le\ & \frac{1}{\epsilon^q} \sum_{j=0}^{n_0-1} \left(\frac{q}{q-1}\right)^q E\Bigg(\sup_{\vartheta\in[u_{n,j},u_{n,j+1}]} |W_{n,\tau,\vartheta} - W_{n,\tau,u_{n,j}}|^q \Bigg).
\end{aligned}
$$

Since $\{W_{n,\tau,\vartheta}, \mathcal{F}_{n,\tau,\vartheta}, \vartheta \ge 0\}$ is a martingale and $\sup_{\vartheta\in[u_{n,j},u_{n,j+1}]} \triangle|W_{n,\tau,\vartheta} - W_{n,\tau,u_{n,j}}| \le \frac{1+\tilde{K}_\tau\tau}{\sqrt{V_n}}$, then following (6.4),

$$
\begin{aligned}
& \frac{1}{\epsilon^q} \sum_{j=0}^{n_0-1} \left(\frac{q}{q-1}\right)^q E\Bigg(\sup_{\vartheta\in[u_{n,j},u_{n,j+1}]} |W_{n,\tau,\vartheta} - W_{n,\tau,u_{n,j}}|^q \Bigg) \\
\le\ & \frac{1}{\epsilon^q} \sum_{j=0}^{n_0-1} \left(\frac{q}{q-1}\right)^q C_q \left(E[\langle W_{n,\tau,u_{n,j}+\cdot}\rangle (u_{n,j+1} - u_{n,j})]^{q/2} + \frac{1+\tilde{K}_\tau\tau}{V_n^{q/2}} \right) \\
\le\ & C_q^*(\epsilon)^{pq/2-p-q} \le \epsilon,
\end{aligned}
$$

where C_q^* is a constant depending only on q and the last inequality holds when ϵ is small enough. Then the desired result follows. □

6.3.2. *Proof of Proposition 6.2.* We need the following lemma (see Lemma 5 in Gu and Lai, 1991).

LEMMA 6.4. *Let $q > 0$ and $r > 1$. Let $\{W_n, n \ge 1\}$ be a sequence of random variables defined in the same probability space and let $\{g_n\}$ be a sequence of nonnegative integrable functions on a measure space $(\mathcal{X}, \mathcal{B}, \mu)$. Suppose that for every fixed $x \in \mathcal{X}$, $g_n(x)$ is nondecreasing in $n \le N$ and that*

$$
E|Z_i - Z_j|^q \le \left(\int_{\mathcal{X}} [g_i(x) - g_j(x)] d\mu(x) \right)^r \text{ for all } 1 \le j \le i \le N.
$$

Then there exists a universal constant $C_{q,r}$ depending only on q and r such that

$$E\left(\sup_{n\le N}|Z_i-Z_j|\right)^q \le C_{q,r}\left(\int_{\mathcal{X}}[g_N(x)-g_1(x)]d\mu(x)\right)^r.$$

PROOF OF PROPOSITION 6.2. Choose positive numbers $p,q>1$ such that $pq/2-p-q>1$. Let $w_0=0$, and define w_j inductively by $w_{j+1}=j\epsilon^p/\tilde{K}_\tau$. Denote $N_0=\left\lfloor \tilde{K}_\tau\tau/\epsilon^p\right\rfloor+1$, and redefine $w_{N_0}=\tau$.

Let $w_{n,i}=iV_n^{-r}$ and $\mathcal{N}_w=\{w_{n,i}:i=0,1,\cdots,\lfloor\tau\cdot V_n^r\rfloor+1\}$. For statement simplicity, assume that V_n takes a constant value. Then

$$P\left(\int_0^\tau\int_{u+w_{n,i}}^{u+w_{n,i+1}}p_n(du\,ds)\ge 2\right)=O(V_n^2)O(V_n^{-2r})=O(V_n^{-2r+2}).$$

From Condition B(iii), it follows that

$$\begin{aligned}(6.5)\qquad & P\left(\sup_{i,w_{n,i}\le w\le w_{n,i+1}}|\tilde{W}_{n,w,\tau}-\tilde{W}_{n,w_{n,i},\tau}|\ge 2V_n^{-1/2}+\tilde{K}_\tau V_n^{-r+1}\right)\\ \le\ & P\left(\sup_i\int_0^\tau\int_{u+w_{n,i}}^{u+w_{n,i+1}}p_n(du\,ds)\ge 2\right)\\ &+P\left(\sup_i\int_0^\tau\int_{u+w_{n,i}}^{u+w_{n,i+1}}q_n(du\,ds)\ge\tilde{K}_\tau V_n^{-r+1+1/2}\right)\\ \le\ & O(V_n^{-r+2})+P(R_\tau\ge V_n^{1+1/2}),\end{aligned}$$

which converges to 0 in probability, according to Condition B(ii), when $r>2$.

Therefore, to prove Proposition 6.2, by (6.5) and the martingale property for $\{\tilde{W}_{n,w,\vartheta},\mathcal{F}_{n,\tau,\vartheta},0<\vartheta\le\tau\}$ along entry time, we only need to show that for any $\epsilon>0$,

$$P\Bigg(\max_{0\le j<N_0}\sup_{\substack{0\le\vartheta\le\tau\\ w\in[w_j,w_{j+1}]\cap\mathcal{N}_w}}|\tilde{W}_{n,w,\vartheta}-\tilde{W}_{n,w_j,\vartheta}|\ge\epsilon\Bigg)\to 0,\quad \textit{as } n\to\infty.$$

For simplicity, we only need to consider the case in which equation (3.1) holds almost surely. Then, by Doob's inequality and (6.4), similar as in the proof of Proposition 6.1,

$$
\begin{aligned}
& P\Bigg(\max_{0\le j<N_0} \sup_{\substack{0\le\vartheta\le\tau \\ w\in[w_j,w_{j+1}]\cap\mathcal{N}_w}} |\tilde{W}_{n,w,\vartheta}-\tilde{W}_{n,w_j,\vartheta}|\ge\epsilon\Bigg) \\
\le\ & \frac{1}{\epsilon^q}\sum_{j=0}^{N_0-1} E\Bigg(\sup_{\substack{0\le\vartheta\le\tau \\ w\in[w_j,w_{j+1}]\cap\mathcal{N}_w}} |\tilde{W}_{n,w,\vartheta}-\tilde{W}_{n,w_j,\vartheta}|^q\Bigg) \\
\le\ & \frac{1}{\epsilon^q}\sum_{j=0}^{N_0-1}\left(\frac{q}{q-1}\right)^q E\Bigg(\sup_{w\in[w_j,w_{j+1}]\cap\mathcal{N}_w} |\tilde{W}_{n,w,\tau}-\tilde{W}_{n,w_j,\tau}|^q\Bigg).
\end{aligned}
$$

Since $\tilde{W}_{n,w_{n,i+1},\vartheta}-\tilde{W}_{n,w_{n_i},\vartheta}$ is a $\{\mathcal{F}_{n,\tau,\vartheta},\vartheta\ge 0\}$ martingale, from (6.4) and (3.1) which holds almost surely now, we have

$$
\begin{aligned}
& E(|\tilde{W}_{n,w_{n,i+1},\tau}-\tilde{W}_{n,w_{n_i},\tau}|^q) \\
\le\ & C_q\left(E[\langle\tilde{W}_{n,w_{n,i+1},\cdot}-\tilde{W}_{n,w_{n_i},\cdot}\rangle(\tau)]^{q/2}+\frac{1}{V_n^{q/2}}\right) \\
\le\ & C\left(\int_0^\tau\Big[\tilde{K}_\tau\cdot 1(x\le w_{n,i+1})-\tilde{K}_\tau\cdot 1(x\le w_{n,i})\Big]\,dx\right)^{q/2},
\end{aligned}
$$

where C is some big constant. Then from Lemma 6.4, there exists $C^*>0$ such that for all large n,

$$
\begin{aligned}
& \frac{1}{\epsilon^q}\sum_{j=0}^{N_0-1}\left(\frac{q}{q-1}\right)^q E\Big(\sup_{w\in[w_j,w_{j+1}]\cap\mathcal{N}_w}|\tilde{W}_{n,w,\tau}-\tilde{W}_{n,w_j,\tau}|^q\Big) \\
\le\ & \frac{1}{\epsilon^q}\sum_{j=0}^{N_0-1}\left(\frac{q}{q-1}\right)^q C\left(\int_0^\tau \tilde{K}_\tau\cdot 1(w_{n,i_{w_j}}<x\le w_{n,i_{w_{j+1}}+1})\,dx\right)^{q/2} \\
\le\ & C^*(2\epsilon)^{pq/2-p-q},
\end{aligned}
\tag{6.6}
$$

where $i_{w_j}=\max\{i: w_{n,i}\le w_j\}$. Thus (6.6) $<\epsilon$ when ϵ is small enough. Then the desired conclusion follows. □

References.

[1] Andersen, P. K., Borgan, Ø., Gill, R. D. and Keiding, N. (1993). *Statistical Models Based on Counting Processes.* Springer.

[2] Arjas, E. and Haara, P. (1984). A marked point process approach to censored failure data with complicated covariates. *Scandinavian Journal of Statistics* **11** 193-209.

[3] Arjas, E. (1989). Survival models and martingale dynamics (with discussion). *Scandinavian Journal of Statistics* **16** 177-225.

[4] Bilias, Y., Gu, M. and Ying, Z. (1997). A general asymptotic theory for cox model with staggered entry. *The Annals of Statistics* **25** 662-682.
[5] Cox, D. R. (1972). Regression Models and Life-Tables. *Journal of the Royal Statistical Society. Series B (Methodological)* **34** 187-220.
[6] Cui, L., Hung, H. M. J. and Wang, S. J. (1999). Modification of sample size in group sequential trials. *Biometrics* **55** 853-857
[7] Doob, J. L. (1953). *Stochastic Processes*. Wiley, New York.
[8] Feng, J. (1999). Estimations in Survival Analysis - A Stochastic Filtering Perspective. *Applied Probability Trust* October 28.
[9] Fisher, L. (1998). Self-designing clinical trials. *Statistics in Medicine* **17** 1551-1562.
[10] Fleming, T. R. and Harrington, D. (1991). *Counting Processes and Survival Analysis*. John Wiley & Sons, Inc.
[11] Friedman, L. M., Furberg, C. D. and DeMets, D. L. (1998). *Fundamentals of Clinical Trials (3rd edition)*. Springer-Verlag, New York.
[12] Gu, M. G. and Lai, T. L. (1991). Weak convergence of time-sequential censored rank statistics with applications to sequential testing in clinical trials. *The Annals of Statistics* **19** 1403-1433.
[13] Hu, F. and Rosenberger, W. F. (2006). *The theory of response-adaptive randomization in clinical trials.* John Wiley & Sons, Inc.
[14] Hu, F., Zhang, L. and He, X. (2009). Efficient randomized-adaptive designs. *The Annals of Statistics* **37** 2543-2560.
[15] Jacod, J. and Shiryaev, A. N. (2003). *Limit Theorems for Stochastic Processes (2nd edition)*. Springer.
[16] Jennison, C. and Turnbull, B. W. (2000). *Group Sequential Methods with Applications to Clinical Trials*. Chapman & Hall/CRC.
[17] Kallianpur, G. and Xiong, J. (1995). *Stochastic Differential Equations in Infinite Dimensional Spaces.* IMS Lecture Notes.
[18] Kalbfleisch, J. D. and Prentice, R. L. (2002). *The Statistical Analysis of Failure Time Data*. New York: Wiley.
[19] Lai, T. L. and Siegmund, D. (1983). Fixed Accuracy Estimation of an Autoregressive Parameter. *The Annals of Statistics* **11** 478-485.
[20] Lan, K. K. G. and DeMets, D. L. (1983). Discrete sequential boundaries for clinical trials. *Biometrika* **70** 659-663.
[21] Last, G. and Brandt, A. (1995). *Marked Point Processes on the Real Line: The Dynamic Approach*. Springer.
[22] Lenglart, E., Lepingle, D. and Pratelli, M. (1980). Presentation unifiee de certaines inégualités de la théorie des martingales. *Séminaire de Probabilités. Lecture Notes in Math* **14** 26-48. Springer, Berlin.
[23] Mantel, N. and Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute* **22** 719-48.
[24] Martinussen, T. and Scheike, T. H. (2000). A nonparametric dynamic additive regression model for longitudinal data. *The Annals of Statistics* **28** 1000-1025.
[25] Martinussen, T. and Scheike, T. H. (2006). *Dynamic Regression Models for Survival Data.* Springer.
[26] Murphy, S. A. and Bingham, D. (2008). Screening Experiments for Developing Dynamic Treatment Regimes. *J. Amer. Statist. Assoc.* **184** 391-408.
[27] O'Brien, P. C. and Fleming, T. R. (1979). a multiple testing procedure for clinical trials. *Biometrics* **35** 549-556.
[28] Pocock, S. J. (1977). Group sequential methods in the design and analysis of clinical

trials. *Biometrika* **64** 191-199.

[29] Pocock, S. J. and Simon, R. (1975). Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. *Biometrics* **31** 103-115.

[30] Proschan, M. A., Lan, K. K. G. and Wittes, J. T. (2006). *Statistical Monitoring of Clinical Trials: A Unified Approach.* Springer.

[31] Rebolledo, R. (1980). Central limit theorem for local martingales. *Z.Wahr. verw. Geb.* **51** 269-286.

[32] Robins, J. M. (1986). A new approach to causal inference in mortality studies with sustained exposure periods-application to control of the healthy worker survivor effect. *Computers and Mathematics with Applications* **14** 1393-1512.

[33] Rosenberger, W. F. and Sverdlov, O. (2008). Handling covariates in the design of clinical trials. *Statistical Science* **23** 404-419.

[34] Sellke, T. and Siegmund, D. (1983). Sequential analysis of the proportional hazards model. *Biometrika* **70** 315-26.

[35] Shen, Y. and Fisher, L. (1999). Statistical inference for self-designing clinical trials with a one-sided hypothesis. *Biometrics* **55** 190-197.

[36] Shen, Y. and Cai, J. (2003). Sample size reestimation for clincal trials with censored survival data. *J. Amer. Statist. Assoc.* **98** 418-426.

[37] Slud, E. (1984). Sequential linear rank tests for two-sample censored survival data. *The Annals of Statistics* **12** 551-571.

[38] Wei, L. J. (1978). The adaptive biased coin design for sequential experiments. *The Annals of Statistics* **6** 92-100.

[39] Wei, L. J. and Durham, S. (1978). The randomized play-the-winner rule in medical trials. *J. Amer. Statist. Assoc.* **73** 840-843.

[40] Zelen, M. (1969). Play the winner and the controlled clinical trial. *J. Amer. Statist. Assoc.* **64** 131-146.

Xiaolong Luo
Biostatistics, Clinical R&D
Celgene Corporation
86 Morris Avenue
Summit, NJ 07901

Gongjun Xu and Zhiliang Ying
Department of Statistics
Columbia University
1255 Amsterdam Avenue
New York, NY 10027
E-mail: gongjun@stat.columbia.edu