

---

# A note on the potential outcomes framework in continuous time

---

**Johan Sebastian Ohlendorff**  
johan.ohlendorff@sund.ku.dk  
University of Copenhagen

## ABSTRACT

In this brief note, we consider the seminal work by [Ryalen \(2024\)](#) and compare it with the approach given in [Rytgaard et al. \(2022\)](#), corresponding to their marked point process settings. We study these works in simple multi-state models.

We consider a multi-state model with states  $\{0, 1, 2, 3\}$  with at most one visitation time for the treatment (that is at most one point where treatment may change), no time-varying covariates, and no baseline covariates. In the initial state everyone starts as treated. We consider the setting with no censoring. The multi-state model is shown in [Figure 1](#).

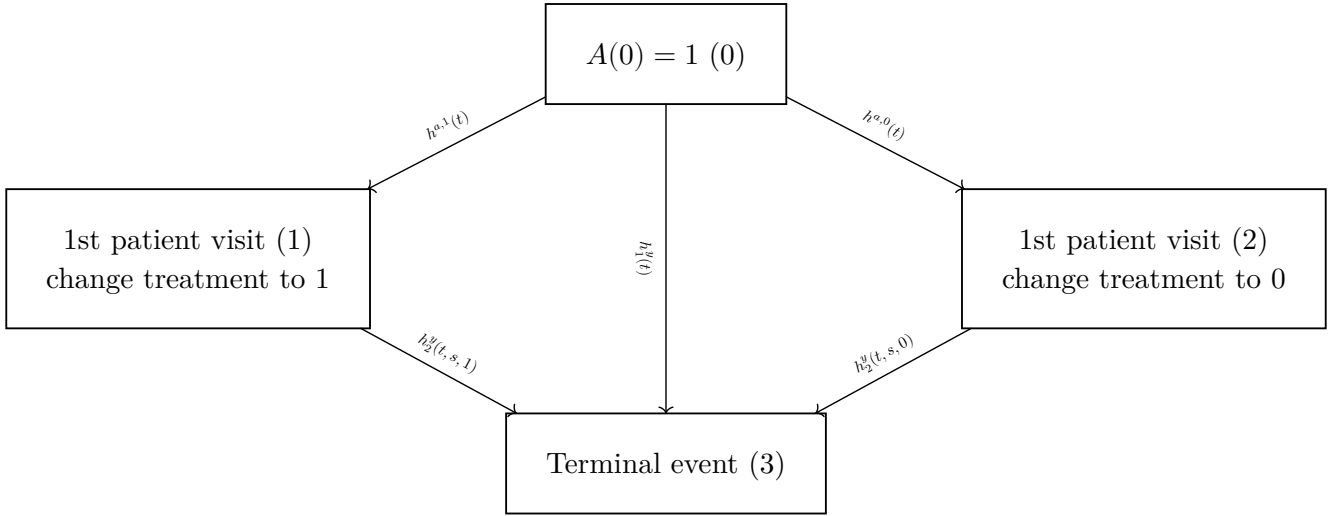


Figure 1: A multi-state model allowing one visitation time for the treatment with the possible treatment values  $\mathcal{A} = \{0, 1\}$ .

The observed times for the jumps between states are

$$\begin{aligned}
T_0 &= \inf\{t > 0 \mid X(t) \neq 0\}, \\
\Delta &= X(T_0), \\
T &= \inf\{t > 0 \mid X(t) = 3\}.
\end{aligned} \tag{1}$$

Also note that  $V = \inf\{t > 0 \mid X(t) \in \{1, 2\}\}$  represents the first visitation time for the treatment. We will intervene on this time, so that the patient stays treated after the first visitation time. but still visits the doctor at the same time.

We can also define the observed counting processes for the multi-state model  $N = (N^{01}, N^{02}, N^{03}, N^{13}, N^{23})$  as

$$\begin{aligned}
N^{01}(t) &= \mathbb{1}\{T_0 \leq t, \Delta = 1\}, \\
N^{02}(t) &= \mathbb{1}\{T_0 \leq t, \Delta = 2\}, \\
N^{03}(t) &= \mathbb{1}\{T_0 \leq t, \Delta = 3\}, \\
N^{13}(t) &= \mathbb{1}\{T \leq t, \Delta = 1\}, \\
N^{23}(t) &= \mathbb{1}\{T \leq t, \Delta = 2\}.
\end{aligned} \tag{2}$$

The corresponding filtration is given by  $\mathcal{F}_t = \sigma(N^{01}(s), N^{02}(s), N^{03}(s), N^{13}(s), N^{23}(s) \mid s \leq t)$ . Additionally the intensities are assumed to be given by,

$$\begin{aligned}
\lambda^{01}(t) &= \mathbb{1}\{t \leq T_0\} h^{a,1}(t), \\
\lambda^{02}(t) &= \mathbb{1}\{t \leq T_0\} h^{a,0}(t), \\
\lambda^{03}(t) &= \mathbb{1}\{t \leq T_0\} h_1^y(t), \\
\lambda^{13}(t) &= \mathbb{1}\{T_0 < t \leq T, \Delta = 1\} h_2^y(t, T_0, 1), \\
\lambda^{23}(t) &= \mathbb{1}\{T_0 < t \leq T, \Delta = 2\} h_2^y(t, T_0, 0),
\end{aligned} \tag{3}$$

in order of the counting processes, that is the compensators are absolutely continuous with respect to the Lebesgue measure (important as we shall see). Here we take  $h^{a,1} : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ ,  $h^{a,0} : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ ,  $h_1^y : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  and  $h_2^y : \mathbb{R}_+ \times \mathbb{R}_+ \times \{0, 1\} \rightarrow \mathbb{R}_+$  to be fixed and non-random functions.

We fit the above multi-state model into the framework of [Ryalen \(2024\)](#) and consider such processes on  $[0, \tau]$  for  $\tau > 0$ . Associated with a marked point process is a counting measure. Let  $(\{0, 1\}, \mathcal{A})$  be the mark space for the treatment. Then let,

$$\begin{aligned}
N_t^a(A) &= N^{01}(t)\mathbb{1}\{1 \in A\} + N^{02}(t)\mathbb{1}\{0 \in A\}, A \subseteq \{0, 1\}, t \in [0, \tau], \\
N_t^y &= N^{03}(t) + N^{13}(t) + N^{23}(t), t \in [0, \tau].
\end{aligned} \tag{4}$$

We consider a multivariate marked point process  $N = (N^a, N^y)$ . Such a process can be considered as a single marked point process, but to make the notation more clear we consider it as a pair of point processes. In [Ryalen \(2024\)](#), the intervention is a mapping, that transforms from the whole marked point process realization to a new treatment process realization. The intervention that we will look at transforms  $N_t^a$  to a new process  $\mathcal{N}_t^a$ , where

$$\mathcal{N}_t^a(A) = g(t, N|_{[0,t]}) = \mathbb{1}\{1 \in A\} \mathcal{N}_t^{01} + 0 = \mathbb{1}\{1 \in A\} \mathcal{N}_t^{01}, \tag{5}$$

for  $g : [0, \tau] \times \mathcal{N}_\tau^{\{0,1\}} \rightarrow \mathbb{R}_+$  where  $\mathcal{N}_\tau^{\{0,1\},d}$  is the space of multivariate marked point processes with  $d$  components. This is the intervention that makes treatment shifts impossible. The intervention is in their terms “optional” but not “predictable” because the intervention  $g$  depends on the information available at time  $t$  and not only on the information strictly prior to  $t$ . Thus the compensating measure of  $\mathcal{N}_t^a$  is given by

$$\mathcal{L}_t^a(A) = \mathbb{1}\{1 \in A\} \int_0^t \lambda^{01}(s) ds = \mathbb{1}\{1 \in A\} \int_0^t \mathbb{1}\{s \leq T_0\} h^{a,1}(s) ds. \tag{6}$$

[Ryalen \(2024\)](#) uses the term canonical compensator to mean that the random part of the compensator only depends on  $N$  and baseline covariates, that is we may write  $\lambda^{ij}(t) = \dot{\lambda}^{ij}(t, N)$  and  $\dot{\lambda}^{i,j}$  is a non-random function. This can be seen, for example for  $\lambda^{13}$ , as

$$\lambda^{13}(t) = \mathbb{1}\{N^{01}(t) = 1, N^{13}(t-) = 0\} h_2^y(t, \inf\{t > 0 \mid N^{01}(t) = 1\}, 1). \tag{7}$$

A multivariate potential outcome counting process  $\tilde{N}$  is a potential outcome process if

1.  $\tilde{N}_t^a$  has compensator  $\mathcal{L}_t^a(\tilde{N})$ .
2.  $\tilde{N}_t^y$  has compensator  $\Lambda^y(t, \tilde{N})$ .

The intuition behind this definition is that the canonical compensators fully determine the distribution of a marked point process. The canonical compensator also makes the marked point processes into a sort-of structural equation system, where we replace the “structural” equations with the compensators.

We consider  $\tau^{A,B} := \inf\{t > 0 \mid N_t^a(B) \neq \mathcal{N}_t^a(B)\}$  for  $B \subseteq \{0,1\}$ . Then  $\tau^A := \inf\{\tau^{A,B} \mid B \subseteq \{0,1\}\}$ . For  $B \subseteq \{0,1\}$  with  $1 \in B$ , we always have  $\mathcal{N}^a(B) = N^a(B)$ . This is trivially also the case with  $B = \emptyset$ . Therefore, we have

$$\tau^A = \inf\{t > 0 \mid N^{02}(t) = 1\}. \quad (8)$$

which is the first time that treatment is changed to 0.

Let  $\mathbb{N}_t = \mathbb{1}\{\tau^A \leq t\}$  and  $\Psi_t^a$  denote the compensator with respect to the filtration  $\mathcal{F}_t$ . If our outcome process is  $Y_t = \mathbb{1}\{T \leq t\}$ , then three basic conditions for identifiability are

1. Consistency:  $\tilde{Y}_t \mathbb{1}\{\tau^A > t\} = Y_t \mathbb{1}\{\tau^A > t\}$ .
2. Exchangeability: The intensity  $\lambda^{12}$  for the filtration  $\mathcal{F}_t$  is the same as it is for  $\mathcal{G}_t = \mathcal{F}_t \vee \sigma(\tilde{Y}_s, \tau \geq s \geq 0)$ .
3. Positivity:  $V_t = \frac{\mathbb{1}\{\tau^A > t\}}{\mathcal{P}_{s \in (0,t]}(1 - d\Psi^a(s))}$  is a uniformly integrable martingale, where  $\mathcal{P}$  denotes the product integral. This is the same as stating that the measure  $Q$  given by  $dQ = V_\tau dP$  is a probability measure.

Then,  $\mathbb{E}[\tilde{Y}_t] = \mathbb{E}[Y_t V_t]$  (Theorem 1 in [Ryalen \(2024\)](#)). Note that  $\mathbb{N}_t = N^{02}(t)$ , so  $\Psi_t^a = \int_0^t \lambda^{02}(s) ds = \int_0^t \mathbb{1}\{s \leq T_0\} h^{a,0}(s) ds$ . This implies that  $V_t = \frac{1 - N_t^{02}}{\exp(-\int_0^t \mathbb{1}\{s \leq T_0\} h^{a,0}(s) ds)}$ . Define  $\mathbb{N}^{a,\psi}(A) = \mathbb{1}\{\tau^A \leq t, \mathcal{N}^a(\{\tau^a\} \times A) = 1\}$ . Then,

$$\begin{aligned} \mathbb{N}_t^{a,\varphi}(A) &:= \mathbb{1}\{\tau^A \leq t, \mathcal{N}^a(\{\tau^a\} \times A) = 1\} \\ &= \mathbb{1}\{\tau^A \leq t\} \mathbb{1}\{\mathcal{N}^a(\{\tau^a\} \times A) = 1\} \\ &= N^{02}(t) (N^{01}(\tau^A) - N^{01}(\tau^A -)) \mathbb{1}\{1 \in A\} \\ &= 0. \end{aligned} \quad (9)$$

The last follows since: If  $N^{02}(t) = 1$ , meaning  $\tau^A \leq t$ , then  $N^{01}(\tau^A) = 0$  and  $N^{01}(\tau^A -) = 0$  because at most one of the processes  $N^{01}$  and  $N^{02}$  can jump. Its compensator  $\mathcal{L}_t^{a,\varphi}(\cdot)$  is therefore 0.

It also follows that  $\Delta\Phi_t^a = 0$  because the compensator is absolutely continuous with respect to the Lebesgue measure. Let  $\mathcal{L}_t^{y,a}$  be the compensator to  $\mathbb{N}_t^{y,a} := \mathbb{1}\{\tau^A \leq t, N^y(\{\tau^a\}) = 1\} = 0$ , i.e., 0. Using Pål’s g-formula (Theorem 2 in [Ryalen \(2024\)](#)), we have

$$\begin{aligned} \mathcal{L}_t^a &= \frac{\mathcal{L}_t^a - \mathcal{L}_t^{a,\varphi}}{1 - \Delta\Phi_t^a} = \mathcal{L}_t^a, \\ \mathcal{L}_t^y &= \frac{\mathcal{L}_t^y - \mathcal{L}_t^{y,a}}{1 - \Delta\Phi_t^a} = \mathcal{L}_t^y. \end{aligned} \quad (10)$$

is the  $Q$ - $\mathcal{F}_t$  compensators of  $N_t^a$  and  $N_t^y$ , respectively. This is the usual g-formula given in [Rytgaard et al. \(2022\)](#). Critically, we used that the compensator is absolutely continuous with respect to the Lebesgue measure. This is not surprising as  $V_\tau$  is actually just the normal inverse probability weights given in [Rytgaard et al. \(2022\)](#). All considerations imply that interventional probability of dying before time  $t$  is given by

$$\begin{aligned}\mathbb{E}[\tilde{Y}_t] &= \int_0^t \exp\left(-\int_0^s h^{a,1}(u) + h_1^y(u) du\right) h_1^y(s) ds \\ &\quad + \int_0^t \exp\left(-\int_0^s h^{a,1}(u) + h_1^y(u) du\right) h^{a,1}(s) \int_s^t \exp\left(-\int_s^w h_2(u, s, 1) du\right) h_2(w, s, 1) dw ds\end{aligned}\tag{11}$$

Now compare this with

$$\begin{aligned}\mathbb{E}[Y_t] &= \int_0^t \exp\left(-\int_0^s h^{a,1}(u) + h^{a,0}(u) + h_1^y(u) du\right) h_1^y(s) ds \\ &\quad + \int_0^t \exp\left(-\int_s^w h^{a,1}(u) + h^{a,0}(u) + h_1^y(u) du\right) h^{a,1}(s) \int_s^t \exp\left(-\int_s^w h_2(u, s, 1) du\right) h_2(w, s, 1) dw ds \\ &\quad + \int_0^t \exp\left(-\int_s^w h^{a,1}(u) + h^{a,0}(u) + h_1^y(u) du\right) h^{a,0}(s) \int_s^t \exp\left(-\int_s^w h_2(u, s, 0) du\right) h_2(w, s, 0) dw ds\end{aligned}$$

which is the observed probability of dying before time  $t$ . According to Pål's product integral formula (Lemma 1 in [Ryalen \(2024\)](#)), the previous equation can be written as

$$\begin{aligned}\mathbb{E}[Y_t] &= \int_0^\infty \exp\left(-\int_0^s h^{a,1}(u) \mathbb{1}\{u \leq \infty\} + h^{a,0}(u) \mathbb{1}\{u \leq \infty\} + h_1^y(u) \mathbb{1}\{u \leq \infty\} du\right) h_1^y(s) \mathbb{1}\{s \leq \infty\} \cdot \mathbb{1}\{s \leq t\} ds \\ &\quad + \int_0^\infty \exp\left(-\int_0^s h^{a,1}(u) \mathbb{1}\{u \leq \infty\} + h^{a,0}(u) \mathbb{1}\{u \leq \infty\} + h_1^y(u) \mathbb{1}\{u \leq \infty\} du\right) h^{a,1}(s) \mathbb{1}\{s \leq \infty\} \cdot 0 ds \\ &\quad + \int_0^\infty \exp\left(-\int_0^s h^{a,1}(u) \mathbb{1}\{u \leq \infty\} + h^{a,0}(u) \mathbb{1}\{u \leq \infty\} + h_1^y(u) \mathbb{1}\{u \leq \infty\} du\right) h^{a,0}(s) \mathbb{1}\{s \leq \infty\} \cdot 0 ds \\ &\quad + \int_0^\infty \exp\left(-\int_0^s h^{a,1}(u) \mathbb{1}\{u \leq \infty\} + h^{a,0}(u) \mathbb{1}\{u \leq \infty\} + h_1^y(u) \mathbb{1}\{u \leq \infty\} du\right) h^{a,1}(s) \mathbb{1}\{s \leq \infty\} \\ &\quad \quad \int_0^\infty \exp\left(-\int_s^w h_2(u, s, 1) \mathbb{1}\{s < u \leq \infty\} du\right) h_2(w, s, 1) \mathbb{1}\{s < w \leq \infty\} \mathbb{1}\{s \leq t\} dw ds \\ &\quad + \int_0^\infty \exp\left(-\int_0^s h^{a,1}(u) \mathbb{1}\{u \leq \infty\} + h^{a,0}(u) \mathbb{1}\{u \leq \infty\} + h_1^y(u) \mathbb{1}\{u \leq \infty\} du\right) h^{a,0}(s) \mathbb{1}\{s \leq \infty\} \\ &\quad \quad \int_0^\infty \exp\left(-\int_s^w h_2(u, s, 0) \mathbb{1}\{s < u \leq \infty\} du\right) h_2(w, s, 0) \mathbb{1}\{s < w \leq \infty\} \mathbb{1}\{s \leq t\} dw ds\end{aligned}\tag{13}$$

Note: The at-risk indicator's are present in the above formulation can be seen by the following: The zero process does not have any jump points. Hence, the first jump point is at infinity in the outer integral. Similarly, in the inner integral, the random measure  $\delta_{t_1, x_1}$  does not have a second jump point.

If we do the calculation corresponding to a multi-state model where one of the intermediate states is treatment discontinuation, corresponding to the deletion of state 1, we will get the ordinary g-formula too.

Extensions could be to stochastic interventions in an observational study with a time to operation. Then it might make sense to intervene on the distribution of the time to operation,

such that its observed hazard ratio is twice as large as the observational one. Currently this does not fit within the framework of [Ryalen \(2024\)](#).

## 1 Multiple event points

Note that

$$\mathbb{1}\{\tau^A \leq t\} = \sum_k \mathbb{1}\{T_k \leq t, D_k = a, A(T_k) = 0, A(T_{k-1}) = \dots = A(T_1) = 1\} \quad (14)$$

In my notes I found that  $\mathbb{1}\{T_k \leq t, D_k = a, A(T_k) = 0\}$  is a martingale with compensator  $\int_0^t \lambda_{s(t,a,1,\mathcal{X})} \mathbb{1}\{T_{k-1} < s \leq T_k\} ds$ .

Consider for  $s < t$

$$\begin{aligned} & \mathbb{E} \left[ \mathbb{1}\{T_k \leq t, D_k = a, A(T_k) = 0, A(T_{k-1}) = \dots = A(T_1) = 1\} \right. \\ & \quad \left. - \int_0^t \lambda_u(a, 1, \mathcal{X}) \mathbb{1}\{T_{k-1} < u \leq T_k\} du \mathbb{1}\{A(T_{k-1}) = \dots = A(T_1) = 1\} \mid \mathcal{F}_s \right] \\ &= \mathbb{E}[\mathbb{1}\{T_{k-1} \leq s\} \left( \mathbb{1}\{T_k \leq t, D_k = a, A(T_k) = 0, A(T_{k-1}) = \dots = A(T_1) = 1\} \right. \\ & \quad \left. - \int_0^t \lambda_u(a, 1, \mathcal{X}) \mathbb{1}\{T_{k-1} < u \leq T_k\} du \mathbb{1}\{A(T_{k-1}) = \dots = A(T_1) = 1\} \mid \mathcal{F}_s \right)] \quad (15) \\ &= \mathbb{E} \left[ \mathbb{1}\{T_k \leq t, D_k = a, A(T_k) = 0\} \right. \\ & \quad \left. - \int_0^t \lambda_s(t, a, 1, \mathcal{X}) \mathbb{1}\{T_{k-1} < s \leq T_k\} ds \mid \mathcal{F}_s \right] \mathbb{1}\{A(T_{k-1}) = \dots = A(T_1) = 1\} \mathbb{1}\{T_{k-1} \leq s\} \\ &= 0 \end{aligned}$$

Since  $\int_0^t \lambda_{s(t,a,1,\mathcal{X})} \mathbb{1}\{T_{k-1} < s \leq T_k, A(T_{k-1}) = \dots = A(T_1) = 1\} ds$  is continuous and adapted, it is predictable. It is also increasing, so it is the compensator to the process  $\mathbb{1}\{T_k \leq t, D_k = a, A(T_k) = 0, A(T_{k-1}) = \dots = A(T_1) = 1\}$ . Thus  $\Delta \Psi_t^a = 0$  (identically zero).

Using this decomposition, we find

$$\begin{aligned} & \mathbb{1}(\tau^A \leq t) \mathbb{1}(\mathcal{N}^a(\{\tau^a\}, \{1\})) \\ &= \sum_k \mathbb{1}\{T_k \leq t, D_k = a, A(T_k) = 0, A(T_{k-1}) = \dots = A(T_1) = 1\} \mathbb{1}(\mathcal{N}^a(\{\tau^a\}, \{1\}) = 1) \\ &= \sum_k \mathbb{1}\{T_k \leq t, D_k = a, A(T_k) = 0, A(T_{k-1}) = \dots = A(T_1) = 1\} (N_{\tau^A}^a(\{1\}) - N_{\tau^A-}^a(\{1\})) \mathbb{1}(A(T_k) = 1) \quad (16) \\ &= 0 \end{aligned}$$

This means that the g-formulas stay the same as with the single event case.

## Bibliography

Ryalen, P. (2024). *On the role of martingales in continuous-time causal inference*.

Rytgaard, H. C., Gerds, T. A., & Laan, M. J. van der. (2022). Continuous-Time Targeted Minimum Loss-Based Estimation of Intervention-Specific Mean Outcomes. *The Annals of Statistics*, 50(5), 2469–2491. <https://doi.org/10.1214/21-AOS2114>