Proposed Algorithms for Time-Dependent Net Benefit under Informative Censoring

Amir Farahmand

Setting and Assumption

We present the proposed algorithms for **model development** and **external validation** of time-dependent net benefit (tNB) for both **binary** and **competing-risk** outcomes under **covariate-dependent informative censoring**. We use inverse probability of censoring weighting (IPCW) under the assumption

$$C \perp\!\!\!\perp T \mid X,$$

where C is the censoring time, T the time-to-event, Z is the observed time which is $Z_i = \min(T_i, C_i)$, δ_i is an indicator for the individual being censored or not on the denoted time point and X the covariates (or a function thereof such as the risk score). Let $\pi_{i,t}$ denote the predicted risk for subject i at time t, and $\widehat{S}_C(u \mid \pi_{i,t})$ the estimated censoring survival at time u conditional on $\pi_{i,t}$.

IPCW Estimates of Se/Sp/prev

The IPCW-based estimators of time-dependent cumulative sensitivity and dynamic specificity (Blanche et al., 2013) are:

$$\widehat{Se}_{IPCW}(z,t) = \frac{\sum_{i=1}^{n} I(\pi_{i,t} > z, Z_i \le t, \delta_i = 1) \ \cdot \ w_i}{\sum_{i=1}^{n} I(Z_i \le t, \delta_i = 1) \ \cdot \ w_i},$$

$$\widehat{Sp}_{IPCW}(z,t) = \frac{\sum_{i=1}^{n} I(\pi_{i,t} \le z, Z_i > t) \ \cdot \ w_i'}{\sum_{i=1}^{n} I(Z_i > t) \ \cdot \ w_i'},$$

where $\widehat{S}_C(\cdot \mid \pi_{i,t})$ is the estimated censoring survival probability conditional on the predicted risk score.

The IPCW event prevalence at t is:

$$\widehat{\text{prev}}_{IPCW}(t) = \frac{1}{n} \sum_{i=1}^{n} w_i \ I(Z_i \le t, \delta_i = 1),$$

with:

$$w_i = \widehat{S}_C(Z_i \mid \pi_{i,t})^{-1}, \qquad w_i' = \widehat{S}_C(t \mid \pi_{i,t})^{-1}.$$

Suppose we have a prediction model that outputs a time–dependent score $\pi_{i,t}$ for subject i at time t. For both **external validation** and **model development**, the net benefit (NB) for a given model and the default *treat-all* strategy at threshold $z \in (0,1)$ can be written as

$$NB_{\text{model}}(z,t) \;=\; \text{prev}(t)\,\text{Se}(z,t) \;-\; \big(1-\text{prev}(t)\big)\,\big(1-\text{Sp}(z,t)\big)\,\frac{z}{1-z}, \tag{1}$$

$$NB_{\text{all}}(z,t) \;=\; \text{prev}(t) \;-\; \big(1-\text{prev}(t)\big)\,\frac{z}{1-z}, \tag{2}$$

where the time–dependent components are defined by

$$\text{prev}(t) = \Pr(T \le t), \qquad \text{Se}(z,t) = \Pr\big(\pi_{i,t} > z \mid T \le t\big), \qquad \text{Sp}(z,t) = \Pr\big(\pi_{i,t} \le z \mid T > t\big),$$

with T the time–to–event and z the decision threshold.
Following the survival-based derivation in *bayesDCA*[1], we can also rewrite the model NB as

$$NB_{\text{model}}(z,t) \;=\; \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}\big(\pi_{i,t} > z\big)\,\left\{\big[1 - S(t \mid \pi_{i,t} > z)\big] - S(t \mid \pi_{i,t} > z)\,\frac{z}{1-z}\right\}, \tag{3}$$

and the treat–all NB as

$$NB_{\text{all}}(z,t) \;=\; \frac{1}{n}\sum_{i=1}^{n} \left\{\big[1 - S(t)\big] - S(t)\,\frac{z}{1-z}\right\}, \tag{4}$$

where $S(t \mid \cdot)$ denotes the (possibly adjusted) survival probability at time t.
To compute *expected* net benefit (ENB) for both development and validation, we adopt a Bayesian approach. In this formulation, we assess the decision rule via the model's selection indicator $\mathbf{1}(\pi_{i,t} > z)$, while the outcome terms inside the braces use the *true risk* $p_{\theta\,i,t}$ (i.e., risk under parameter draw θ) to propagate parameter uncertainty:

$$NB_{\text{model}}(z,t;\theta) \;=\; \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}\big(\pi_{i,t} > z\big)\,\left\{\big[1 - S(t \mid p_{\theta\,i,t} > z)\big] - S(t \mid p_{\theta\,i,t} > z)\,\frac{z}{1-z}\right\}. \tag{5}$$

$$NB_{\text{all}}(z,t;\theta) \;=\; \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}\big(\pi_{i,t} > z\big)\,\left\{\big[1 - S(t)\big] - S(t)\,\frac{z}{1-z}\right\}. \tag{6}$$

Averaging over the posterior $P(\theta \mid \mathcal{D})$ yields

$$\overline{NB}_{\text{model}}(z,t) \;=\; \mathbb{E}_{\theta\sim P(\theta\mid\mathcal{D})}\big[\,NB_{\text{model}}(z,t;\theta)\,\big], \quad \overline{NB}_{\text{all}}(z,t) \;=\; \mathbb{E}_{\theta\sim P(\theta\mid\mathcal{D})}\big[\,NB_{\text{all}}(z,t;\theta)\,\big],$$

which we approximate via posterior simulation (e.g., Bayesian bootstrap).

Under this formulation, the conditional survival terms in equation (5) can be expressed in terms of $\text{prev}(t), \text{Se}(z,t)$ and $\text{Sp}(z,t)$ as follows:

$$S(t \mid p_{\theta\,i,t} > z) \;=\; \frac{(1-\text{prev}(t))\,(1-\text{Sp}(z,t))}{\Pr(p_{\theta\,i,t} > z)}, \tag{7}$$

$$1 - S(t \mid p_{\theta\,i,t} > z) \;=\; \frac{\text{prev}(t)\,\text{Se}(z,t)}{\Pr(p_{\theta\,i,t} > z)}. \tag{8}$$

These identities allow us to connect the survival-based form of NB directly to the time-dependent $\text{prev}(t), \text{Se}(z,t)$, and $\text{Sp}(z,t)$ when the true risks $p_{\theta\,i,t}$ are used.
we replace the predicted scores $\pi_{i,t}$ with the true risks $p_{\theta\,i,t}$ in the definitions of sensitivity and specificity:

$$\text{Se}(z,t) \;=\; \Pr\!\big(p_{\theta\,i,t} > z \mid T \le t\big), \qquad \text{Sp}(z,t) \;=\; \Pr\!\big(p_{\theta\,i,t} \le z \mid T > t\big).$$

External Validation

Proposed Algorithm

Algorithm 1 External Validation: IPCW tNB under Informative Censoring

Require: Fixed threshold $z \in (0, 1)$; fixed time $t > 0$; validation dataset $\mathbf{v}$ with size n; number of Monte Carlo simulations R; prediction model M; censoring model C.

Ensure: Estimated average net benefit $\text{ENB}^V_{\text{model}}(z, t)$.

Using the proposed prediction model M, **calculate the predicted risks** for each individual in the validation sample: compute $\pi_{i,t}$ for all $i \in \{1, 2, \ldots, n\}$.

for $b = 1, \ldots, R$ **do**

Draw $\mathcal{D}^{*(b)}$ a **Bayesian bootstrap** of the data $\mathbf{v}$ which is a random draw from the posterior distribution of the population given the sample. This can be done via assigning a vector of weights to each observation, with weights coming from Dirichlet$(1, \ldots, 1)$.

Using the proposed model M to calculate predicted risks (true risk) $p^{(b)}_{\theta_{i,t}}$ from $\mathcal{D}^{*(b)}$

Fit censoring model: Using C as the Censoring model and $\mathcal{D}^{*(b)}$, estimate $\widehat{S}^{(b)}_C(u \mid p^{(b)}_{\theta_{i,t}})$ for $u \in \{Z_i, t\}$.

Compute IPCW weights: For each i,

$$w_i^{(b)} = \frac{1}{\widehat{S}^{(b)}_C(Z_i \mid p^{(b)}_{\theta_{i,t}})} \qquad w_i'^{(b)} = \frac{1}{\widehat{S}^{(b)}_C(t \mid p^{(b)}_{\theta_{i,t}})}.$$

Estimate time-dependent components: Using the appropriate formulas compute

$$\widehat{\text{prev}}^{(b)}(t), \quad \widehat{Se}^{(b)}(z, t), \quad \widehat{Sp}^{(b)}(z, t)$$

where the event and non-event groups are weighted by $w_i^{(b)}$ and $w_i'^{(b)}$.

Calculate $\widehat{S}\Big(t \mid p^{(b)}_{\theta_{i,t}} > z\Big)$ **using** $\widehat{\text{prev}}^{(b)}(t)$, $\widehat{Se}^{(b)}(z, t)$, $\widehat{Sp}^{(b)}(z, t)$.

Compute the NB_{model} by plugging $\widehat{S}\Big(t \mid p^{(b)}_{\theta_{i,t}} > z\Big)$ into equation (5).

end for

Aggregate:

$$\text{ENB}^V_{\text{model}}(z, t) = \frac{1}{R} \sum_{b=1}^{R} \text{NB}^{(b)}_{model}(z, t).$$

Bias Assessment

Algorithm 2 Bias Assessment for External Validation: IPCW tNB under Informative Censoring via Simulation Data

Require: Fixed threshold $z \in (0,1)$; fixed time $t > 0$; number of Monte Carlo simulations R; recommended large sample size N and a validation sample size n; prediction model M; censoring model C.

Ensure: Estimated bias $\widehat{\text{Bias}}^{V}_{\text{model}}(z,t)$ and a percentile CI.

1: **Generate oracle (large) validation dataset:** Draw a dataset $\mathcal{V}_{\text{large}}$ of size N from the data-generating mechanism (treated as our population); calculate the predicted risk $(\pi^{\star}_{i,t})$ for all $i \in \{1,2,\ldots,N\}$.

2: **Compute oracle NB (with no informative censoring):** Using known event times compute the *no informative censoring* time-dependent components

$$\text{prev}^{\star}(t), \quad Se^{\star}(z,t), \quad Sp^{\star}(z,t),$$

by simple empirical proportions (i.e., without IPCW/KM). Define

$$\text{NB}^{\star}(z,t) \;=\; \text{prev}^{\star}(t)\, Se^{\star}(z,t) \;-\; \big(1-\text{prev}^{\star}(t)\big)\,\big(1-Sp^{\star}(z,t)\big)\,\frac{z}{1-z}.$$

3: **Generate a dataset** $\mathcal{V}$ with size n from the data-generating mechanism.

4: **Perform Algorithm 1 excluding step 11** on $\mathcal{V}$ to obtain the vector of Net Benefits $\{NB^{(1)}_{model}(z,t), NB^{(2)}_{model}(z,t), \ldots, NB^{(R)}_{model}(z,t)\}$.

5: **Aggregate (Bias and CI):**

$$\Delta_b = NB^{(b)}_{model}(z,t) - NB^{\star}(z,t), \qquad \widehat{\text{Bias}}^{\text{V}}_{\text{model}}(z,t) = \frac{1}{R}\sum_{b=1}^{R}\Delta_b.$$

A $(1-\alpha)$ percentile CI for the bias is given by

$$\big[\text{Quantile}_{\alpha/2}(\Delta_b),\ \text{Quantile}_{1-\alpha/2}(\Delta_b)\big].$$

EVPI

Algorithm 3 EVPI for External Validation: IPCW tNB under Informative Censoring

Require: Fixed threshold $z \in (0,1)$; fixed time $t > 0$; validation dataset $\mathbf{v}$ with size n; number of Monte Carlo simulations R; prediction model M; censoring model C.

Ensure: Estimated $EVPI^V(z,t)$.

1: Using the proposed prediction model, **calculate the predicted risks** for each individual in the validation sample: compute $\pi_{i,t}$ for all $i \in \{1,2,\ldots,n\}$.

2: **for** $b = 1,\ldots,R$ **do**

3: Draw $\mathcal{D}^{*(b)}$ a **Bayesian bootstrap** of the data $\mathbf{v}$ which is a random draw from the posterior distribution of the population given the sample. This can be done via assigning a vector of weights to each observation, with weights coming from Dirichlet$(1,\ldots,1)$.

4: **Using the proposed model** M to calculate predicted risks $p^{(b)}_{\theta_{i,t}}$ from $\mathcal{D}^{*(b)}$

5: **Fit censoring model:** Using C as the Censoring model and $\mathcal{D}^{*(b)}$, estimate $\widehat{S}^{(b)}_C(u \mid p^{(b)}_{\theta_{i,t}})$ for $u \in \{Z_i, t\}$.

6: **Compute IPCW weights:** For each i,

$$w_i^{(b)} = \frac{1}{\widehat{S}^{(b)}_C(Z_i \mid p^{(b)}_{\theta_{i,t}})} \qquad w_i'^{(b)} = \frac{1}{\widehat{S}^{(b)}_C(t \mid p^{(b)}_{\theta_{i,t}})}.$$

7: **Estimate time-dependent components:** Using the appropriate formulas compute

$$\widehat{\text{prev}}^{(b)}(t), \quad \widehat{Se}^{(b)}(z,t), \quad \widehat{Sp}^{(b)}(z,t)$$

where the event and non-event groups are weighted by $w_i^{(b)}$ and $w_i'^{(b)}$.

8: **Calculate** $\widehat{S}\Big(t \mid p^{(b)}_{\theta_{i,t}} > z\Big)$ and $\widehat{S}(t)$ **using** $\widehat{\text{prev}}^{(b)}(t)$, $\widehat{Se}^{(b)}(z,t)$, $\widehat{Sp}^{(b)}(z,t)$.

9: **Compute the** NB_{model} **and** NB_{all} in draw b by plugging $\widehat{S}\Big(t \mid p^{(b)}_{\theta_{i,t}} > z\Big)$ and $\widehat{S}(t)$ into equation (5) and (6)

10: **end for**

11: **Aggregate:**

$$ENB^V_{\text{model}}(z,t) = \frac{1}{R}\sum_{b=1}^{R} NB^{(b)}_{\text{model}}(z,t), \quad ENB^V_{\text{all}}(z,t) = \frac{1}{R}\sum_{b=1}^{R} NB^{(b)}_{\text{all}}(z,t),$$

$$\overline{NB}^V_{\max}(z,t) = \frac{1}{R}\sum_{b=1}^{R} \max\!\Big(0,\, NB^{(b)}_{\text{all}}(z,t),\, NB^{(b)}_{\text{model}}(z,t)\Big).$$

12: **Compute EVPI:**

$$EVPI^V(z,t) \;=\; \overline{NB}^V_{\max}(z,t) \;-\; \max\big\{\,0,\, ENB^V_{\text{model}}(z,t),\, ENB^V_{\text{all}}(z,t)\big\}\,.$$

Model Development

Proposed Algorithm

Algorithm 4 Model Development: IPCW tNB under Informative Censoring

Require: Fixed threshold $z \in (0,1)$; fixed time $t > 0$; development dataset **d** with size n; number of Monte Carlo simulations R; prediction model M; censoring model C.

Ensure: Estimated average net benefit $\mathrm{ENB}^D_{\mathrm{model}}(z,t)$.

1: Using the proposed prediction model, **generate the predicted risks** for each individual in the development sample: compute $\pi_{i,t}$ for all $i \in \{1, 2, \dots, n\}$.

2: **for** $b = 1, \dots, R$ **do**

3: Draw $\mathcal{D}^{*(b)}$ a **Bayesian bootstrap** of the data **d** which is a random draw from the posterior distribution of the population given the sample. This can be done via assigning a vector of weights to each observation, with weights coming from Dirichlet$(1, \dots, 1)$.

4: **Fit prediction model on $\mathcal{D}^{*(b)}$:** Train the chosen prediction model M and compute individual risks $p^{(b)}_{\theta_{i,t}}$ (**true risks**) for all $i \in \{1, 2, \dots, n\}$.

5: **Fit censoring model:** Using $\mathcal{D}^{*(b)}$, estimate $\widehat{S}^{(b)}_C(u \mid p^{(b)}_{\theta_{i,t}})$ for $u \in \{Z_i, t\}$.

6: **Compute IPCW weights:** For each i,

$$w_i^{(b)} = \frac{1}{\widehat{S}^{(b)}_C(Z_i \mid p^{(b)}_{\theta_{i,t}})} \qquad w_i'^{(b)} = \frac{1}{\widehat{S}^{(b)}_C(t \mid p^{(b)}_{\theta_{i,t}})}.$$

7: **Estimate time-dependent components:** Using the appropriate formulas compute

$$\widehat{\mathrm{prev}}^{(b)}(t), \quad \widehat{Se}^{(b)}(z,t), \quad \widehat{Sp}^{(b)}(z,t),$$

where the event and non-event groups are weighted by $w_i^{(b)}$ and $w_i'^{(b)}$.

8: **Calculate** $\widehat{S}\Big(t \mid p^{(b)}_{\theta_{i,t}} > z\Big)$ **using** $\widehat{\mathrm{prev}}^{(b)}(t)$, $\widehat{Se}^{(b)}(z,t)$, $\widehat{Sp}^{(b)}(z,t)$.

9: **Compute the** NB_{model} by plugging $\widehat{S}\Big(t \mid p^{(b)}_{\theta_{i,t}} > z\Big)$ into equation (5).

10: **end for**

11: **Aggregate:**

$$\mathrm{ENB}^D_{\mathrm{model}}(z,t) = \frac{1}{R}\sum_{b=1}^{R} \mathrm{NB}^{(b)}_{model}(z,t).$$

Bias Assessment

Algorithm 5 Bias Assessment for Model Development: IPCW tNB under Informative Censoring via Simulation Data

Require: Fixed threshold $z \in (0, 1)$; fixed time $t > 0$; number of Monte Carlo simulations R; recommended large sample size N and a development sample size n; prediction model M; censoring model C.

Ensure: Estimated bias $\widehat{\text{Bias}}^{D}_{\text{model}}(z, t)$ and a percentile CI.

1: **Generate oracle (large) development dataset:** Draw a dataset $\mathcal{D}_{\text{large}}$ of size N from the data-generating mechanism (treated as population); fit a prediction model on $\mathcal{D}_{\text{large}}$ and compute risks $\pi^{\star}_{i,t}$ for all $i \in \{1, 2, \dots, N\}$.

2: **Compute oracle NB (with no informative censoring):** Using known event times compute the *no informative censoring* time-dependent components

$$\text{prev}^{\star}(t), \quad Se^{\star}(z, t), \quad Sp^{\star}(z, t),$$

by simple empirical proportions (i.e., without IPCW/KM). Define

$$\text{NB}^{\star}(z, t) \;=\; \text{prev}^{\star}(t)\, Se^{\star}(z, t) \;-\; \big(1 - \text{prev}^{\star}(t)\big)\, \big(1 - Sp^{\star}(z, t)\big)\, \frac{z}{1 - z}.$$

3: **Generate a dataset** $\mathcal{D}$ with size n from the data-generating mechanism.

4: **Perform Algorithm 4 excluding step 11** on $\mathcal{D}$ to obtain the vector of Net Benefits $\{NB^{(1)}(z, t), NB^{(2)}(z, t), \dots, NB^{(R)}(z, t)\}$.

5: **Aggregate (Bias and CI):**

$$\Delta_b = NB^{(b)}_{model}(z, t) - NB^{\star}(z, t), \qquad \widehat{\text{Bias}}^{\text{D}}_{\text{model}}(z, t) = \frac{1}{R}\sum_{b=1}^{R} \Delta_b.$$

A $(1 - \alpha)$ percentile CI for the bias is given by

$$\big[\text{Quantile}_{\alpha/2}(\Delta_b),\ \text{Quantile}_{1-\alpha/2}(\Delta_b)\big].$$

For the EVPI calculation in the development setting, we also require the *oracle optimal model* net benefit, where the true risk $p_{\theta\, i,t}$ is used both inside the outcome terms and in the decision indicator:

$$NB_{\max}(z, t; \theta) \;=\; \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}\big(p_{\theta\, i,t} > z\big)\, \left\{ \big[1 - S(t \mid p_{\theta\, i,t} > z)\big] - S(t \mid p_{\theta\, i,t} > z)\, \frac{z}{1 - z} \right\}. \tag{9}$$

This expression corresponds to the **true model-based strategy**, where both the selection rule and the outcome are evaluated under the true risk. It provides the upper bound against which the observed model and the treat-all strategy are compared in computing EVPI.

EVPI

Algorithm 6 EVPI for Model Development: IPCW tNB under Informative Censoring

Require: Fixed threshold $z \in (0,1)$; fixed time $t > 0$; development dataset $\mathbf{d}$ with size n; number of Monte Carlo simulations R;prediction model M; censoring model C.

Ensure: Estimated $EVPI^D(z,t)$.

1: Using the proposed prediction model, **generate the predicted risks** for each individual in the development sample: compute $\pi_{i,t}$ for all $i \in \{1,2,\dots,n\}$.

2: **for** $b = 1,\dots,R$ **do**

3: Draw $\mathcal{D}^{*(b)}$ a **Bayesian bootstrap** of the data $\mathbf{d}$ which is a random draw from the posterior distribution of the population given the sample. This can be done via assigning a vector of weights to each observation, with weights coming from Dirichlet$(1,\dots,1)$.

4: **Fit prediction model M on $\mathcal{D}^{*(b)}$:** Train the chosen prediction model and compute individual risks $p^{(b)}_{\theta_{i,t}}$ for all $i \in \{1,2,\dots,n\}$.

5: **Fit censoring model:** Using C as the Censoring model and $\mathcal{D}^{*(b)}$, estimate $\widehat{S}^{(b)}_C(u \mid p^{(b)}_{\theta_{i,t}})$ for $u \in \{Z_i, t\}$.

6: **Compute IPCW weights:** For each i,

$$w_i^{(b)} = \frac{1}{\widehat{S}^{(b)}_C(Z_i \mid p^{(b)}_{\theta_{i,t}})} \qquad w_i'^{(b)} = \frac{1}{\widehat{S}^{(b)}_C(t \mid p^{(b)}_{\theta_{i,t}})}.$$

7: **Estimate time-dependent components:** With appropriate weighted formulas, compute

$$\widehat{\text{prev}}^{(b)}(t), \quad \widehat{Se}^{(b)}(z,t), \quad \widehat{Sp}^{(b)}(z,t)$$

8: **Calculate** $\widehat{S}\Big(t \mid p^{(b)}_{\theta_{i,t}} > z\Big)$ **using** $\widehat{\text{prev}}^{(b)}(t)$, $\widehat{Se}^{(b)}(z,t)$, $\widehat{Sp}^{(b)}(z,t)$.

9: **Compute the NB_{model}, NB_{all} and NB_{max}** in draw b by plugging $\widehat{S}\Big(t \mid p^{(b)}_{\theta_{i,t}} > z\Big)$ and $\widehat{S}(t)$ into equation (5), (6) and (9)

10: **end for**

11: **Aggregate:**

$$ENB^D_{\text{model}}(z,t) = \frac{1}{R}\sum_{b=1}^{R} NB^{(b)}_{\text{model}}(z,t), \quad ENB^D_{\text{all}}(z,t) = \frac{1}{R}\sum_{b=1}^{R} NB^{(b)}_{\text{all}}(z,t),$$

$$\overline{NB}^D_{\max}(z,t) = \frac{1}{R}\sum_{b=1}^{R} NB^{(b)}_{\max}(z,t).$$

12: **Compute EVPI:**

$$EVPI^D(z,t) \;=\; \overline{NB}^D_{\max}(z,t) \;-\; \max\big\{\,0,\; ENB^D_{\text{model}}(z,t),\; ENB^D_{\text{all}}(z,t)\big\}\,.$$

Notes.

- In general, the number of iterations should be high enough such that the Monte Carlo standard error around ENB and EVPI (development or validation) is small compared to its point estimate.
- In both Bias Assessment algorithms, the "oracle" $\mathrm{NB}^{\star}(z,t)$ uses known event times and *no informative censoring adjustment*; it is a large-sample reference (finite N_{large}).
- Using the NB vector we can also report the Variance and the Mean-Squared-Error.

Formulation and Notation

For both **development** and **validation** we consider the time–dependent net benefit (tNB) framework. Let $\pi_{i,t}$ denote the predicted risk score for subject i at time t, and $p_{\theta\, i,t}$ the *true risk* under model parameter θ. The event time is T, the censoring time C, and the observed time $Z = \min(T, C)$. Treatment at threshold $z \in (0,1)$ is defined via the indicator $\mathbf{1}(\cdot)$.

Prevalence, sensitivity, and specificity. We define the time–dependent prevalence, sensitivity, and specificity as

$$\mathrm{prev}(t) = \Pr(T \le t), \quad \mathrm{Se}(z,t) = \Pr(\pi_{i,t} > z \mid T \le t), \quad \mathrm{Sp}(z,t) = \Pr(\pi_{i,t} \le z \mid T > t).$$

These quantities can also be written in terms of $p_{\theta\, i,t}$ when the true risk is available.

Development. For **development**, our formulation is the same for EVPI and EVSI [4, 5]. Both are written in Bayesian terms: the decision indicator $\mathbf{1}(\pi_{i,t} > z)$ comes from the model, while the outcome terms inside the net benefit formulas are evaluated using the true risk $p_{\theta\, i,t}$. This Bayesian formulation allows us to propagate parameter uncertainty and properly assess the expected net benefit.

Validation. For **validation**, the existing formulations of EVPI and EVSI [2, 3] are not expressed in Bayesian terms, but are based directly on estimators of sensitivity, specificity, and prevalence. we adapt this notation and express validation in the same unified framework.

Problem. Our main concern arises in the **validation** setting. In the current literature, both EVSI and EVPI formulations for validation [2, 3] are not expressed in Bayesian terms. EVSI Validation directly use the net benefit formulation in equation (1), without explicitly calculating $\pi_{i,t}$ [3] and then assessing performance against the true risk and also the EVPI Validation use the hard label for calculating the NB and not using the true risk [2]. This creates an inconsistency: while the development formulations are Bayesian and incorporate the true risk, the validation formulations do not. Since our aim is to present all these results together in a single framework, we must adopt consistent notation across both development and validation.

References

[1] Giuliano N. F. Cruz and Keegan Korthauer. Bayesian decision curve analysis with bayesdca, 2023.

[2] Mohsen Sadatsafavi, Tae Y. Lee, Laure Wynants, Andrew J. Vickers, and Paul Gustafson. Value-of-information analysis for external validation of risk prediction models. *Medical Decision Making*, 43(5):564–575, Jul 2023. Epub 2023 Jun 22.

[3] Mohsen Sadatsafavi, Andrew J. Vickers, Tae Y. Lee, Paul Gustafson, and Laure Wynants. Expected value of sample information calculations for risk prediction model validation. *Medical Decision Making*, 45(3):232–244, Apr 2025. Epub 2025 Feb 18.

[4] Mohsen Sadatsafavi, Tae Yoon Lee, and Paul Gustafson. Uncertainty and the value of information in risk prediction modeling. *Medical Decision Making*, 42(5):661–671, Jul 2022. Epub 2022 Feb 25.

[5] Abdollah Safari, Paul Gustafson, and Mohsen Sadatsafavi. Expected value of sample information calculations for risk prediction model development, 2025.