# BUDGET-CONSTRAINED ACTIVE LEARNING TO DECENSOR SURVIVAL DATA

**Anonymous authors**
Paper under double-blind review

## ABSTRACT

Standard supervised learners attempt to learn a model from a **labeled** dataset. Given a small set of labeled instances, and a pool of unlabeled instances, a **budgeted learner** can use its given budget to pay to acquire the labels of some unlabeled instances, which it can then use to produce a model. Here, we explore budgeted learning in the context of survival datasets, which include (right) censored instances, where we know only a lower bound on an instance's time-to-event. Here, that learner can pay to (partially) label a censored instance – *e.g.*, to acquire the actual time for an instance [perhaps go from (3 yr, censored) to (7.2 yr, uncensored)], or other variants [*e.g.*, learn about one more year, so go from (3 yr, censored) to either (4 yr, censored) or perhaps (3.2 yr, uncensored)]. This serves as a model of real world data collection, where follow-up with censored patients does not always lead to uncensoring, and how much information is given to the learner model during data collection is a function of the budget and the nature of the data itself. We provide both experimental and theoretical results for how to apply state-of-the-art budgeted learning algorithms to survival data and the respective limitations that exist in doing so. For an entropy-capped batch objective, marginal greedy selection under equal costs achieves the same $(1 - 1/e)$ approximation factor as BatchBALD. Moreover, empirical analyses on several survival tasks show that our model performs better than other potential approaches on several benchmarks. Our acquisition rule, $\mathrm{EPIG}_{\mathrm{surv}}$, selects a patient and a follow-up depth by the expected reduction in predictive uncertainty on a target population per unit of expected cost. Probes are acquired in batches, with the model refitted between rounds.

## 1 INTRODUCTION

The success of a learner in building an effective model is fundamentally reliant on the quality of the data to which it has access. Data significantly influences the bias, variance, and generalizability of any model produced by a learner (Domingos, 2012). However, data availability is often constrained by several factors, including the cost of labeling the data instances and the need for sufficient diversity in the labeled data to ensure effective generalization to the target distribution. The "*data selection*" challenge is deciding which unlabeled instances to label (from a pool of unlabeled instances), to build an effective model. For example, given a dataset with 50 labeled instances and a budget to acquire labels for 10 additional unlabeled instances from a large pool of unlabeled data, the goal is to leverage the initial 50 labeled instances to identify the 10 whose labels will best augment the first 50 labels, to produce the most accurate prediction model.

This paper extends this technology to the domain of survival prediction, which involves time-to-event data, which is essential in various fields as the event could be failure, recovery, or death, for an instance. In a clinical example, a breast cancer study, researchers could monitor 300 cancer patients over a 5-year span to measure the time until death from any cause. The goal here is to develop a model to predict the time of death for a patient diagnosed with that cancer. Similarly, in a financial study, the objective might be to produce a model that could predict the time until a stock crashes. Unlike standard regression tasks, survival models are often trained and evaluated on data where the time-to-event label is not always precisely observed; instead, some instances are *right-censored – i.e.*, only a lower bound on their event time is observed (Kleinbaum & Klein, 2012).



Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

Censoring is common in real-world datasets – *e.g.*, in the above medical example, some patients might leave the study early (and so be lost to follow-up), while others might be alive when the study is completed. It is difficult to train (and evaluate) a model using such censored data, as most supervised learning approaches handle only fully labeled data. One naïve approach is to simply exclude these censored instances when building the model – however, this can remove a large amount of information from the learning system (more than $70\%$ of instances can be censored) and may result in a biased model – note this will typically remove the long-term survivors (Qi et al., 2024; Turkson et al., 2021).

The field of *Active Learning (AL)* aims to identify which unlabeled instances to label, to construct the dataset that can produce the best model. Its goal is to minimize the amount of additional training data needed to achieve a target accuracy (Ren et al., 2021). Some AL studies compare the number of queries required for training to reach a convergence criterion. Such comparisons depend on the stopping rule. For example, convergence is commonly declared when the loss changes by less than a small value, $\epsilon$, over a predefined number of time steps referred to as "*patience*" (Ren et al., 2021). This imprecision arises because both $\epsilon$ and *patience* are chosen heuristically, with their values varying between tasks and models. As a result, the definition of true convergence remains unclear and inconsistent.

The field of *Budgeted Learning (BL)* addresses this issue by focusing on selecting data instances *within a budget* to minimize model loss, thus making the acquisition resource constraint explicit (Lizotte et al., 2003). BL augments AL by giving the learner a specific overall budget $B \in \mathbb{R}_{>0}$, before it begins data selection, to spend to obtain labels. A good budgeted learner should perform well at **any** given budget (rather than only at large budgets or by converging faster). Here, we apply BL to real-world survival datasets, where the budgeted learner can pay to (partially) "de-censor" a censored instance by determining the status of that instance $k$-years in the future: *e.g.*, learn about $k$ = 1 more year, so go from (3yr, censored) to perhaps (3.2yr, uncensored) or (4yr, censored). We refer to this $k$ as the "probe depth", which is an input to our BL system. If the oracle has access to an instance's event time, a probe of depth $k = \infty$ reveals it.

In some cases, the cost of each $k$-year probe is the same for all instances – which we call the *uniform setting*. We will also consider the non-uniform situation, where there may be different costs for obtaining $k$-year information about different instances.

Although there are now many methods to learn models from survival data (Kleinbaum & Klein, 2012; Cox, 1972; Yu et al., 2011; Haider et al., 2020), very few methods have been developed for data selection with survival data (Vinzamuri et al., 2014), and we are not aware of any published method in this setting that combines budget constraints, a choice of probe depth, and costs that depend on the follow-up actually disclosed. This paper addresses that gap. In our realised-cost settings, the charge depends on the follow-up observed before the probe ends (Appendix A.8). We ask which patient to follow, and for how long, when additional follow-up may reveal an event, end in natural censoring, or reach the requested depth without either endpoint.

**Significance of our Work:** The settings explored in this work cover a wide range of applications; (1) Consider a follow-up 5-year clinical study on some of the patients in the study mentioned above. This is equivalent to having an oracle that informs for $k = 5$ more years of the time to event for any participant in the study. (2) Industrial reliability settings explore the time until failure (Meeker & Escobar, 1998) – *e.g.*, testing time until a light bulb stops working – where a cost may be attached to testing for an additional $k$ days. (3) Many computational problems can be framed as survival problems where the event in question is the time an algorithm will finish (Eggensperger et al., 2018). Running the algorithm for an additional $k$ minutes costs a given price for this computation. Every one of the settings above can be imagined with unique instance costs (different light bulbs may cost different amounts to test or different algorithm inputs may require different compute for the same amount of time). The resource purchased is follow-up, not a guaranteed event label: even a probe that reveals no event consumes budget and changes the observed survival record. The useful decision is therefore which follow-up improves population prediction per unit cost, rather than simply which query reveals the most events. Appendix A.3 further discusses the difference between the terms used in this paper (such as BL, AL, SA, etc.) as well as how our work contributes to these areas of study.

**Our contributions.** (1) We formulate budgeted follow-up acquisition with uniform, random and realised costs, and with either a fixed probe depth or a depth menu. (2) We develop $\mathrm{EPIG}_{\mathrm{surv}}$,

which values the response that follow-up can actually reveal, predicts its cost, and permits repeated probes based on observed history. An entropy-capped surrogate connects batch selection to weighted maximum coverage; marginal greedy selection achieves a $(1 - 1/e)$ approximation under equal costs. (3) We show why bounded-lookahead target-focused policies lack an unconditional whole-run approximation guarantee, and derive one under bounded growth (assumption A1). (4) Against seven baselines on three survival datasets, our rule wins all 21 paired AULC comparisons in each fixed-depth setting; variable-depth, fresh-seed, second-learner and synthetic experiments test the scope of this advantage.

## 2 Related Work

Active learning includes uncertainty sampling (Liu & Li, 2023), committee disagreement (Burbidge et al., 2007) and diversity-based selection (Moses et al., 2023). Batch acquisition limits refitting costs by collecting several observations before updating the model (Ren et al., 2021).

### 2.1 Active learning for survival analysis

Cox-based approaches (Nezhad et al., 2019; 2018; Vinzamuri et al., 2014) select survival observations but do not address our combination of partial follow-up, depth selection and realised costs. C-BALD (Hüttel et al., 2025) measures parameter information in an observed value and its censoring indicator; our adaptation applies it to a probe's response. BatchBALD (Kirsch et al., 2019) accounts for shared parameter information across a batch. IDEAL (Bemporad, 2023) combines predictive uncertainty with distance-based exploration.

The closest follow-up framework is that of Dedja et al. (2023): a query reveals a random fraction of an instance's remaining recorded follow-up, and the instance can be queried again. Their oracle determines the disclosed amount; the acquisition rule does not choose depth or allocate a budget measured in follow-up cost. We include their uncertainty-based query strategy among the baselines and give every score-based comparator the same expected-cost divisor (Appendix D).

**Relation to EPIG.** Expected predictive information gain (Bickford Smith et al., 2023), building on prediction-oriented experimental design (MacKay, 1992), targets predictions rather than parameters. We retain that objective. Its full-label, equal-cost formulation does not specify the likelihood of a finite follow-up response, how to predict outcome-dependent cost, or how to select a patient and depth while permitting repeated probes. These are necessary parts of our survival formulation (Section 4). A further difficulty is that joint prediction-targeted information need not be submodular (Krause & Guestrin, 2005). Our entropy-capped surrogate recovers a batch approximation guarantee, while the sequential analysis identifies both an impossibility result and a sufficient condition for whole-run performance.

## 3 Formulation of the problem

The learner receives a budget $B \in \mathbb{R}_{>0}$ and a dataset $D = \{[x_i, t_i, \delta_i, c_i]\}_{i=1}^{L}$. Here $x_i$ denotes covariates, $t_i$ the observed time, and $\delta_i$ the event indicator: $\delta_i = 1$ if the event is observed and 0 if the observation is right-censored. The event time $T_i$ is latent when $\delta_i = 0$. Known costs are denoted by $c_i$; under realised follow-up costs, the learner must predict a cost for each proposed action.

An action $a = (i, k)$ extends patient $i$'s follow-up by depth $k > 0$, either fixed or chosen from a menu. For example, a five-year probe of $(2.4, \text{censored})$ can return an event at 3.5, censoring at 7.4, or earlier natural censoring. Actions are acquired in batches of at most $b$ distinct patients, with refitting between rounds; unresolved patients may be probed again. The total realised cost must not exceed $B$, and performance is evaluated throughout the budget.

The model has parameters $\omega \sim p(\omega \mid D)$ and maps covariates to probabilities over disjoint event-time bins $\{b_j\}_{j=1}^{n}$. The latent label $Y_i$ identifies the bin containing $T_i$, not the bin containing a censoring time. An uncensored record reveals $Y_i$; a censored record restricts the possible event times. Tail sums of the bin probabilities define an individual survival distribution $S(t \mid x, \omega) = P(T > t \mid x, \omega)$ (Haider et al., 2020).

We assume independent censoring, $T \perp C \mid X$, for the survival likelihood and external censored negative log-likelihood (NLL) (Rindt et al., 2022; Lillelund et al., 2025). Our primary budget-wide metric is the area under the external NLL learning curve (AULC). Secondary metrics are the IPCW integrated Brier score (IBS) (Graf et al., 1999), the C-index and mean absolute error with pseudo-observations (MAE-PO) (Qi et al., 2023a); MAE-PO evaluates predicted event times in days while retaining censored test patients (Appendix J).

# 4 Our Method

The survival setting changes both what can be learned and what must be paid. Information about an unobservable event-time tail cannot justify a short probe, and a longer requested window may terminate early. $\mathrm{EPIG}_{\mathrm{surv}}$ therefore combines a finite-window response likelihood with posterior-predicted cost over patient–depth actions. It uses hypothetical fresh outcomes at target covariates, permits repeated probes based on observed history, and supplies an entropy-capped objective for batch analysis.

## 4.1 Adjusting the predictive distribution for censoring

Let $j_i$ denote the bin containing patient $i$'s current censoring time. The discrete learner removes mass from earlier bins and renormalises:

$$p_{\mathrm{cens}}(Y_i = b_r \mid x_i, \omega, D) = \frac{p(Y_i = b_r \mid x_i, \omega, D)}{\sum_{\ell=j_i}^{n} p(Y_i = b_\ell \mid x_i, \omega, D)}, \qquad r \geq j_i. \tag{1}$$

This is a bin-level approximation: it retains the bin containing $t_i$. The response model then allocates event mass over the portion of each bin intersected by the probe window. All information-based methods use the same adjusted probabilities.

## 4.2 The oracle's response

For candidate $i$, let $t_i$ be its current censoring time, $T_i$ its event time, and $k_i$ its probe depth. The acquisition must score what the oracle will actually return, rather than the unobservable full event time. In the basic two-branch oracle this response is

$$Z_i := \begin{cases} (T_i, 1), & t_i < T_i \leq t_i + k_i, \\ (t_i + k_i, 0), & T_i > t_i + k_i, \end{cases} \tag{2}$$

where the second branch is a later-censored observation. If follow-up can terminate earlier for another reason, $Z_i$ instead records that observed censoring time and status; the result below is unchanged provided the probability model is for this complete response. This distinction matters empirically: ordinary event-time MI and variance can credit posterior disagreement in the unobservable tail beyond the probe window, whereas $Z_i$ collapses that tail to the single response the oracle returns.

Our implementation uses three branches — an event inside the window, a natural censoring inside it, and "no endpoint by $t_i + k_i$" — with event mass distributed uniformly within each bin intersected by the window; the natural-censoring branch and the expected-cost model use one reverse Kaplan–Meier estimate shared across patients (covariate-free), whereas the likelihood requires only $T \perp C \mid X$. Every response-based score below is computed from these per-draw response probabilities $p(z_i \mid \hat{\omega}_m)$ of the $K$-sample ensemble.

## 4.3 Population-targeted response information: $\mathrm{EPIG}_{\mathrm{surv}}$

**Target information and the evaluated rule.** Let $X_V$ be target covariates representing deployment; we use 300 pool patients selected without their outcomes. At each $x_v$, $Y_v$ is a hypothetical fresh outcome, conditionally independent of a probe response given $\omega$. The Bayes risk under logarithmic loss is predictive entropy, so

$$H(Y_V \mid X_V, D) - \mathbb{E}_{Z_S \mid D} H(Y_V \mid X_V, Z_S, D) = I(Y_V; Z_S \mid X_V, D). \tag{3}$$

For a sum of marginal target log losses, a single action has value

$$f_a(D) = \sum_{v \in V} I(Y_v; Z_a \mid x_v, D). \tag{4}$$

We estimate it by reweighting the posterior ensemble with each possible response likelihood and measuring the reduction in target entropy. With expected action cost $\hat{c}_a$ (Appendix A.8), the evaluated rule is

$$\text{rank actions by } f_a(D)/\hat{c}_a \text{ and select up to } b \text{ budget-feasible actions.} \tag{5}$$

At most one action per patient is accepted in a batch (Appendix Algorithm 3). With equal costs, this is top-$b$ selection; at $b = 1$ it is one-step Bayes-optimal for the model's marginal target log-loss. This identity does not guarantee the realised held-out improvement of an approximately refitted model.

For the parameter-information comparison and sequential model, we assume mutual conditional independence of responses from distinct patients:

$$p(z_S \mid \omega, D) = \prod_{a \in S} p(z_a \mid \omega, D). \tag{6}$$

This makes parameter information submodular (Kirsch et al., 2019), but does not make prediction-targeted joint information submodular.

**An entropy-capped batch objective.** Selecting a batch by the sum of single-probe scores, $\sum_{a \in S} f_a$, can double-count information shared by probes and assign a target more information than its entropy, although $I(Y_v; Z_S \mid x_v, D) \le H(Y_v \mid x_v, D)$ always. The corresponding joint-response objective $F_V(S) := \sum_v I(Y_v; Z_S \mid x_v, D)$ accounts for dependence between probes but is not submodular even under Equation (6) — two probes can be worthless alone and decisive together (Appendix A.7) — so the standard submodular greedy guarantee does not apply. We therefore keep the sum but cap each target's credit at its entropy,

$$C(S) \;=\; \sum_{v \in V} \min\Big\{ \sum_{a \in S} I(Y_v; Z_a \mid x_v, D),\; H(Y_v \mid x_v, D) \Big\}. \tag{7}$$

The cap gives batch selection a *coverage-like* structure: each target is an element of weight $h_v = H(Y_v \mid x_v, D)$, probe $a$ covers an amount $I(Y_v; Z_a \mid x_v, D)$ of it, and a fully covered element yields nothing more. Weighted maximum coverage — every probe covers each element fully or not at all — is the special case (Theorem A.4), giving the hardness reduction in Appendix A.5.1. As with BatchBALD, the greedy approximation guarantee follows from submodularity. For equal costs and one candidate action per patient (fixed depth), feasibility is a cardinality constraint. The objective $C$ is normalised, monotone and submodular and agrees with $f_a$ on single probes (Proposition A.3); greedy selection attains $C(S_{\text{greedy}}) \ge (1 - 1/e) \max_{|S| \le b} C(S)$ (Nemhauser et al., 1978); and no polynomial-time algorithm can guarantee a larger constant for this objective over general information values unless $\mathrm{P} = \mathrm{NP}$ (Theorem A.4, Feige, 1998). With non-uniform costs and variable depth the rule ranks by gain per expected cost, a density heuristic without this guarantee (Section 4.4).

**Connection to the evaluated batches.** The submodularity of $C$ follows from its capped non-negative sums and does not require conditional independence. Its guarantee applies to marginal greedy selection. For the evaluated top-$b$ batch $S_{\text{top}}$, a stronger conclusion holds whenever no target's cap binds on that batch: writing $M(S) = \sum_{a \in S} f_a$, every feasible $S$ satisfies $C(S) \le M(S) \le M(S_{\text{top}}) = C(S_{\text{top}})$. Thus that batch maximises $C$ exactly. This certificate has not been verified at every real-data round (Appendix A.7).

### 4.4 Budgeted coverage and sequential guarantees

With known non-uniform costs, maximising $C$ under $\sum_{a \in S} c_a \le B$ is a submodular knapsack problem containing budgeted maximum coverage. Gain per cost is motivated by coverage algorithms; partial enumeration recovers $(1 - 1/e)$ for budgeted coverage (Khuller et al., 1999) and for monotone submodular knapsack (Sviridenko, 2004). The evaluated rule uses *singleton* gain per expected cost, while marginal greedy uses $C(S \cup \{a\}) - C(S)$ (Algorithm 1). They coincide when the relevant caps do not bind. We claim no such approximation guarantee for the evaluated heterogeneous-cost or depth-menu heuristic. For realised costs, a probe incurs $c_0 + \min(T_i, C_i^{\text{nat}}, t_i + k) - t_i$, where $c_0$ is overhead. Feasibility reserves the upper charge $c_0 + k$ per action; expected cost is used only for ranking, so realised cumulative spending stays within $B$.

**Across rounds.** A parity construction shows that policies restricted to target–response joint laws of at most $L$ probes can achieve an arbitrarily small fraction of the optimal adaptive value: $L + 1$ individually uninformative probes may jointly reveal the target (Theorem B.1). This rules out an

unconditional approximation guarantee for that policy class, not optimal planning in general. For equal-cost probes and exact posterior updates, let $d(a \mid h)$ be marginal target information at history $h$ and $m_h = \max_a d(a \mid h)$. Assumption **A1** requires $d(a \mid h') \leq \kappa m_h$ after any extension $h'$ by at most $K$ observations. With $K$ probes in $R = K/b$ rounds, top-$b$ or capped greedy then satisfies

$$\mathcal{V}(\pi) \geq \left[1 - \left(1 - \frac{1}{\kappa K}\right)^R\right] \mathrm{OPT}_K. \tag{8}$$

Here $\mathcal{V}$ is expected whole-run target information. A1 is sufficient, not necessary; we do not certify it for the fitted learners (Appendix B).

### 4.5 Bayesian Downstream Learner

We use Bayesian multi-task logistic regression (MTLR) (Yu et al., 2011; Qi et al., 2023b), which handles censored observations without imputation and does not assume proportional hazards. Four bootstrap variational fits with sixteen weight draws each form a 64-sample ensemble; refits are warm-started after every acquisition round. Every method receives this same ensemble. The likelihood, prior discussion and estimators appear in Appendix A.2; a random survival forest provides a second-learner comparison (Appendix H).

## 5 Experiments

We compare seven baselines: C-BALD, BatchBALD, IDEAL, the query strategy of Dedja et al., Entropy, Variance and Random (Appendix D). All use the same learner and refitting schedule; every score-based rule divides by the same expected-cost estimate when costs are non-uniform. Random is an uninformed reference baseline. Uniform-cost experiments additionally include a refit-only control and a label-informed oracle reference (Appendix K).

**Data and protocol.** We use SUPPORT (9,105 patients, 42 features, 32% censored) (Knaus et al., 1995), FLCHAIN (7,874; 8; 72%) (Dispenzieri et al., 2012), and MIMIC-IV (38,319; 91; 74%) (Johnson et al., 2023; Qi et al., 2023a). Per seed, 3,000 patients form the development set and up to 4,000 disjoint patients form the test set. In fixed-depth runs, 160 historical records retain their observed follow-up, with 100 reserved for validation; the remaining development patients enter training with independent artificial censoring $C_i^{\mathrm{art}} \sim U(0, q_{0.5})$, unless a recorded endpoint occurs first. All real-data batches contain at most ten patients. Appendix E gives the splits, censoring construction, cost models and depth quantiles.

**Four settings.** Write $k_q$ for the $q$-quantile of residual time among historical records. With fixed depth $k_{0.25}$, uniform costs allocate 400 probes; independent random prices $c_i \sim U(0.2, 0.8)$ allocate 200 cost units; realised costs charge overhead plus disclosed follow-up, with budget $575k_{0.25}$. These settings respectively isolate patient selection, price sensitivity and outcome-dependent cost. The variable-depth setting offers $\{k_{0.1}, k_{0.25}, k_{0.5}\}$ under realised costs and budget $200k_{0.25}$, testing the trade-off between many short and fewer long probes. Its initial-censoring protocol also differs, so it is not an isolated depth ablation.

**Statistics.** All comparisons are paired by seed. AULC is the area of the external NLL curve over the budget axis divided by its length; with random or realised follow-up costs the curves are evaluated on the common horizon reached by every (seed, method) pair. For each comparator we report the mean paired difference against ours with a 20,000-resample bootstrap 95% interval, classified W (interval above zero), L (below) or T (contains zero); T means *unresolved at this* $n$, not equivalence. The three fixed-depth settings use 40 development seeds; the uniform-cost setting is additionally confirmed on 40 fresh seeds on MIMIC-IV and SUPPORT, and the variable-depth setting uses 40 fresh pre-registered seeds per dataset.

## 6 Results

Table 1 summarises the main result: $\mathrm{EPIG}_{\mathrm{surv}}$ wins all 21 paired AULC comparisons in each of the three fixed-depth settings, and 19 of 21 with variable depth, with two unresolved comparisons and no resolved losses. This is evidence across the tested budgets and cost regimes; it does not imply that every baseline is worse on every metric or on every individual seed.

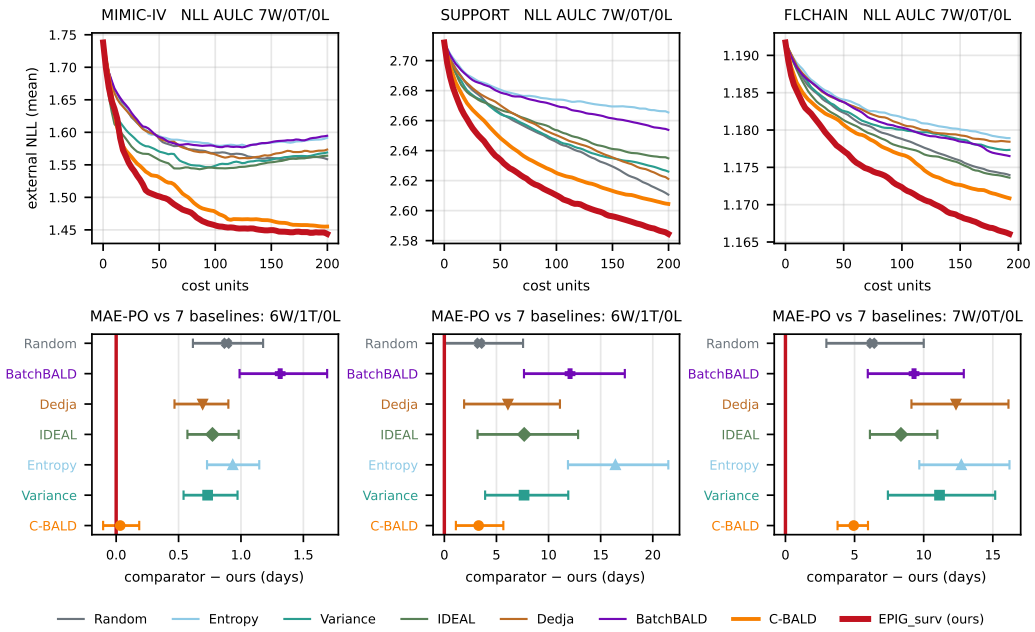



Figure 1: Random costs, all three datasets. Top: external censored NLL against cumulative cost on the common horizon, mean over seeds; ours is the thick red line; panel titles give the paired AULC record against the seven baselines. Bottom: MAE-PO at the final budget, each comparator minus ours in days, paired by seed, with 95% bootstrap intervals; right of the red line means ours is better. $n = 40$ seeds.

Table 1: Paired AULC wins/ties/losses against seven baselines per dataset. Each fixed-depth setting uses 40 development seeds; variable depth uses 40 fresh seeds. W/T/L are defined in Section 5.

| Setting | MIMIC-IV | SUPPORT | FLCHAIN |
|---|---|---|---|
| Uniform costs | 7W/0T/0L | 7W/0T/0L | 7W/0T/0L |
| Random costs | 7W/0T/0L | 7W/0T/0L | 7W/0T/0L |
| Realised costs, fixed depth | 7W/0T/0L | 7W/0T/0L | 7W/0T/0L |
| Variable depth | 7W/0T/0L | 6W/1T/0L | 6W/1T/0L |

### 6.1 Uniform costs: one depth, equal cost

Table 2 gives the mean external NLL of every method at four budgets: $\mathrm{EPIG}_{\mathrm{surv}}$ is lowest in 12 of the 12 columns, by at least 0.0006. Paired over the 40 seeds it is 7W/0T/0L on MIMIC-IV, 7W/0T/0L on SUPPORT and 7W/0T/0L on FLCHAIN on AULC against the seven baselines (learning curves and the paired record in Appendix K). Against C-BALD, the mean AULC improvements are $+0.0131$, $+0.0022$ and $+0.0013$ on MIMIC-IV, SUPPORT and FLCHAIN, respectively. On 40 fresh seeds with the protocol unchanged, MIMIC-IV and SUPPORT give 6W/1T/0L and 6W/1T/0L on AULC (Appendix Table 20); the ties are C-BALD on MIMIC-IV and Random on SUPPORT. Under a joint Holm correction across all 21 pilot comparisons 21 survive. On the point-prediction metric MAE-PO (Appendix J) $\mathrm{EPIG}_{\mathrm{surv}}$ is 6W/1T/0L on MIMIC-IV, 3W/4T/0L on SUPPORT and 6W/1T/0L on FLCHAIN at 400 probes. On SUPPORT, halving or doubling the number of time bins preserves the advantage over Random, while $\mathrm{EPIG}_{\mathrm{surv}}$ and C-BALD remain close (Appendix I).

Two dataset-specific mechanisms behind these numbers — re-probing on MIMIC-IV and a label-informed oracle that violates ignorability on FLCHAIN — are analysed in Appendix K.

Table 2: Uniform costs, 40 development seeds: mean external censored NLL (lower is better) of every method at four budgets, with its standard error over the seeds beneath; best value per column in bold.

| | MIMIC-IV | | | | SUPPORT | | | | FLCHAIN | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| probes | 100 | 200 | 300 | 400 | 100 | 200 | 300 | 400 | 100 | 200 | 300 | 400 |
| $\mathrm{EPIG}_{\mathrm{surv}}$ (ours) | **1.507** ±0.023 | **1.464** ±0.016 | **1.450** ±0.013 | **1.448** ±0.014 | **2.646** ±0.008 | **2.624** ±0.009 | **2.610** ±0.009 | **2.599** ±0.009 | **1.180** ±0.003 | **1.175** ±0.003 | **1.171** ±0.003 | **1.168** ±0.003 |
| C-BALD | 1.532 ±0.025 | 1.479 ±0.016 | 1.463 ±0.014 | 1.455 ±0.013 | 2.649 ±0.008 | 2.625 ±0.009 | 2.613 ±0.009 | 2.604 ±0.010 | 1.181 ±0.003 | 1.177 ±0.003 | 1.173 ±0.003 | 1.170 ±0.003 |
| Entropy | 1.595 ±0.020 | 1.580 ±0.016 | 1.585 ±0.016 | 1.590 ±0.016 | 2.681 ±0.009 | 2.674 ±0.008 | 2.669 ±0.008 | 2.666 ±0.008 | 1.184 ±0.003 | 1.182 ±0.003 | 1.180 ±0.003 | 1.179 ±0.003 |
| Variance | 1.569 ±0.023 | 1.551 ±0.018 | 1.557 ±0.017 | 1.568 ±0.018 | 2.664 ±0.009 | 2.646 ±0.010 | 2.635 ±0.010 | 2.626 ±0.011 | 1.183 ±0.003 | 1.180 ±0.003 | 1.179 ±0.003 | 1.177 ±0.003 |
| IDEAL | 1.556 ±0.020 | 1.545 ±0.015 | 1.553 ±0.015 | 1.564 ±0.015 | 2.667 ±0.008 | 2.653 ±0.009 | 2.641 ±0.010 | 2.635 ±0.011 | 1.181 ±0.003 | 1.178 ±0.003 | 1.175 ±0.003 | 1.173 ±0.003 |
| Dedja | 1.590 ±0.018 | 1.560 ±0.015 | 1.565 ±0.015 | 1.570 ±0.015 | 2.671 ±0.008 | 2.651 ±0.010 | 2.635 ±0.010 | 2.622 ±0.011 | 1.184 ±0.003 | 1.181 ±0.003 | 1.179 ±0.003 | 1.178 ±0.003 |
| BatchBALD | 1.593 ±0.018 | 1.579 ±0.016 | 1.584 ±0.016 | 1.592 ±0.017 | 2.679 ±0.009 | 2.670 ±0.009 | 2.661 ±0.009 | 2.654 ±0.010 | 1.184 ±0.003 | 1.180 ±0.003 | 1.178 ±0.003 | 1.176 ±0.003 |
| Random | 1.590 ±0.019 | 1.568 ±0.015 | 1.560 ±0.014 | 1.561 ±0.016 | 2.665 ±0.007 | 2.646 ±0.008 | 2.629 ±0.009 | 2.611 ±0.009 | 1.182 ±0.003 | 1.179 ±0.003 | 1.176 ±0.003 | 1.173 ±0.003 |

### 6.2 Random costs: one depth, instance-dependent costs

Figure 1 shows the learning curves: on SUPPORT and FLCHAIN no baseline's mean curve falls below ours at any budget, and on MIMIC-IV only within the first 15 of 200 cost units. Paired over seeds, $\mathrm{EPIG}_{\mathrm{surv}}$/cost is 7W/0T/0L on MIMIC-IV, 7W/0T/0L on SUPPORT and 7W/0T/0L on FLCHAIN on AULC against the seven baselines (Appendix Table 19). On MAE-PO it is 6W/1T/0L on MIMIC-IV, 6W/1T/0L on SUPPORT and 7W/0T/0L on FLCHAIN (Figure 1, bottom; Appendix J).

### 6.3 Extensions: realised costs and a menu of depths

The same rule extends without change to the two further settings, reported in full in the appendix. Charging each probe its realised follow-up at one depth (Appendix F) leaves the record intact: 7W/0T/0L on MIMIC-IV, 7W/0T/0L on SUPPORT and 7W/0T/0L on FLCHAIN against the seven baselines on AULC. Offering a menu of three depths under realised cost (variable depth; Appendix G), confirmed on 40 fresh pre-registered seeds per dataset with candidate, roster and endpoints frozen beforehand, gives no resolved loss to any baseline on NLL, AULC or IBS over 63 paired comparisons (56 wins, 7 unresolved).

**Second learner.** Replacing the MTLR by a random survival forest, with every rule, the protocol and the metrics unchanged, leaves $\mathrm{EPIG}_{\mathrm{surv}}$ without a resolved AULC loss on SUPPORT or FLCHAIN; on MIMIC-IV, short probes produce little change in that learner and the comparison is inconclusive (Appendix H).

**Synthetic studies.** With exact posterior updates over a finite hypothesis set and exact evaluation, $\mathrm{EPIG}_{\mathrm{surv}}$ has no resolved AULC loss to the 7 baselines in seven pre-registered worlds (22W/27T/0L), beats all of them in a world built from the assumptions of Section 4.3 at each tested batch size (1, 5, 10 and 20), and has resolved losses in a world with a menu of depths (4W/1T/2L; Appendix C).

## 7 Conclusion

This paper introduced a generalised form of active learning that incorporates budget constraints, individual instance costs, and the partial-information structure of survival data. The acquisition rule values each finite-window response by its expected reduction in predictive uncertainty on the target population, divided by its expected cost. Across the three datasets, the fixed-depth experiments show no statistically resolved AULC loss against the reported baselines; the fresh-seed uniform-cost comparisons support this finding. Final-budget MAE-PO likewise shows no resolved loss (48W/15T/0L over 63 development comparisons). The variable-depth confirmation yields no resolved

loss on NLL, AULC or IBS over 63 comparisons. With a random survival forest, the AULC results on SUPPORT and FLCHAIN provide further evidence across learners (Appendix H). Under equal costs, marginal greedy selection on the entropy-capped objective achieves a $(1 - 1/e)$ approximation, optimal for the general objective class unless $\mathrm{P} = \mathrm{NP}$. The evaluated top-$b$ batch maximises this surrogate whenever no cap binds on that batch. A separate whole-run guarantee holds under bounded growth.

Four limitations scope the claims of this work and are natural targets for future study: (1) The batch guarantee concerns marginal greedy selection on the entropy-capped surrogate $C$ under equal costs. It does not bound the true joint-response objective or automatically apply to top-$b$ selection. The whole-run bound additionally requires A1, which we do not certify on fitted models. (2) Maximizing model-internal information need not minimize held-out predictive NLL; we evaluate this relationship empirically. (3) The $(1 - 1/e)$ approximation statement applies only to equal costs; the non-uniform density rule is a heuristic. (4) The main results use one Bayesian linear MTLR with an unconverged initial fit; the forest replication yields informative comparisons on two datasets, while short probes produce little change on the third (Appendix H). The response channel is undefined for draws with no mass after a candidate's follow-up (7 rows on MIMIC-IV; Appendix K).

We anticipate that this emerging intersection of budgeted active learning and survival analysis will motivate many further studies, of which the four targets above are only the most immediate.

### Reproducibility statement

The numerical result tables, experimental figures and generated numerical macros are produced from recorded per-seed experiment files by the accompanying generation scripts; the acquisition harness was frozen (hash recorded) before the confirmation seeds ran, and all seed files, invariant checks and reports are archived with the code. Appendix E specifies the experimental protocol; Appendices A.7 and B provide the batch and sequential proofs.

### AI use statement

Large language models were used for editing, for drafting and checking proofs and counterexamples, and for code review; every result, proof and number was verified by the authors against the recorded data and the code.

## References

Per Kragh Andersen and Maja Pohar Perme. Pseudo-observations in survival analysis. *Statistical Methods in Medical Research*, 19(1):71–99, 2010.

Per Kragh Andersen, John P. Klein, and Susanne Rosthøj. Generalised linear models for correlated pseudo-observations, with applications to multi-state models. *Biometrika*, 90(1):15–27, 2003.

Alberto Bemporad. Active learning for regression by inverse distance weighting. *Information Sciences*, 626:275–292, 2023.

Freddie Bickford Smith, Andreas Kirsch, Sebastian Farquhar, Yarin Gal, Adam Foster, and Tom Rainforth. Prediction-oriented Bayesian active learning. In *Proceedings of the 26th International Conference on Artificial Intelligence and Statistics (AISTATS)*, volume 206 of *PMLR*, pp. 7331–7348, 2023.

Robert Burbidge, Jem J. Rowland, and Ross D. King. Active learning for regression based on query by committee. In *Intelligent Data Engineering and Automated Learning (IDEAL)*, volume 4881 of *LNCS*, pp. 209–218, 2007.

David R. Cox. Regression models and life-tables. *Journal of the Royal Statistical Society: Series B*, 34(2):187–220, 1972.

Klest Dedja, Felipe Kenji Nakano, and Celine Vens. Active learning for survival analysis with incrementally disclosed label information. In *Proceedings of the Workshop on Interactive Adaptive Learning (IAL@ECML-PKDD 2023)*, volume 3470 of *CEUR Workshop Proceedings*, pp. 46–64, 2023.

Angela Dispenzieri, Jerry A. Katzmann, Robert A. Kyle, et al. Use of nonclonal serum immunoglobulin free light chains to predict overall survival in the general population. *Mayo Clinic Proceedings*, 87(6):517–523, 2012.

Pedro Domingos. A few useful things to know about machine learning. *Communications of the ACM*, 55(10):78–87, 2012.

Katharina Eggensperger, Marius Lindauer, and Frank Hutter. Neural networks for predicting algorithm runtime distributions. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 1442–1448, 2018. RECONSTRUCTED: stands in for the original manuscript's run-time-as-survival citation.

Uriel Feige. A threshold of $\ln n$ for approximating set cover. *Journal of the ACM*, 45(4):634–652, 1998.

Erika Graf, Claudia Schmoor, Willi Sauerbrei, and Martin Schumacher. Assessment and comparison of prognostic classification schemes for survival data. *Statistics in Medicine*, 18(17–18):2529–2545, 1999.

Humza Haider, Bret Hoehn, Sarah Davis, and Russell Greiner. Effective ways to build and evaluate individual survival distributions. *Journal of Machine Learning Research*, 21(85):1–63, 2020.

Neil Houlsby, Ferenc Huszár, Zoubin Ghahramani, and Máté Lengyel. Bayesian active learning for classification and preference learning. arXiv:1112.5745, 2011.

Frederik Boe Hüttel, Christoffer Riis, Filipe Rodrigues, and Francisco Camara Pereira. Bayesian active learning for censored regression. In *Machine Learning and Knowledge Discovery in Databases (ECML PKDD 2025)*, volume 16014 of *LNCS*, pp. 36–51, 2025.

Alistair E. W. Johnson, Lucas Bulgarelli, Lu Shen, et al. MIMIC-IV, a freely accessible electronic health record dataset. *Scientific Data*, 10:1, 2023.

Samir Khuller, Anna Moss, and Joseph (Seffi) Naor. The budgeted maximum coverage problem. *Information Processing Letters*, 70(1):39–45, 1999.

Andreas Kirsch, Joost van Amersfoort, and Yarin Gal. BatchBALD: efficient and diverse batch acquisition for deep Bayesian active learning. In *Advances in Neural Information Processing Systems 32 (NeurIPS)*, 2019.

David G. Kleinbaum and Mitchel Klein. *Survival Analysis: A Self-Learning Text*. Springer, 3rd edition, 2012.

William A. Knaus, Frank E. Harrell, Joanne Lynn, et al. The SUPPORT prognostic model: objective estimates of survival for seriously ill hospitalized adults. *Annals of Internal Medicine*, 122(3): 191–203, 1995.

Andreas Krause and Carlos Guestrin. Near-optimal nonmyopic value of information in graphical models. In *Proceedings of the 21st Conference on Uncertainty in Artificial Intelligence (UAI)*, 2005.

Christian Marius Lillelund, Shi-ang Qi, Russell Greiner, and Christian Fischer Pedersen. Position: Stop chasing the C-index when evaluating survival analysis models. *arXiv preprint arXiv:2506.02075*, 2025.

Shang Liu and Xiaocheng Li. Understanding uncertainty sampling via equivalent loss. arXiv:2307.02719, 2023.

Daniel J. Lizotte, Omid Madani, and Russell Greiner. Budgeted learning of naive-Bayes classifiers. In *Proceedings of the 19th Conference on Uncertainty in Artificial Intelligence (UAI)*, 2003.

David J. C. MacKay. Information-based objective functions for active data selection. *Neural Computation*, 4(4):590–604, 1992.

William Q. Meeker and Luis A. Escobar. *Statistical Methods for Reliability Data*. Wiley, 1998.

Aimee Moses, Christopher W. Bogart, Kelsey O'Haire, Lidia Solorzano, and Elinor Yeo. Diversity-based active learning: creating a representative object detection dataset in 3D point clouds. In *Geospatial Informatics XIII (Proc. SPIE 12525)*, 2023. doi: 10.1117/12.2663179.

George L. Nemhauser, Laurence A. Wolsey, and Marshall L. Fisher. An analysis of approximations for maximizing submodular set functions—I. *Mathematical Programming*, 14:265–294, 1978.

Milad Zafar Nezhad, Najibesadat Sadati, Kai Yang, and Dongxiao Zhu. A deep active survival analysis approach for precision treatment recommendations: application of prostate cancer. arXiv:1804.03280, 2018. RECONSTRUCTED: original target unknown; not cited in the new paper.

Milad Zafar Nezhad, Najibesadat Sadati, Kai Yang, and Dongxiao Zhu. A deep active survival analysis approach for precision treatment recommendations: application of prostate cancer. *Expert Systems with Applications*, 115:16–26, 2019. doi: 10.1016/j.eswa.2018.07.070.

Shi-ang Qi, Neeraj Kumar, Mahtab Farrokh, Weijie Sun, Li-Hao Kuan, Rajesh Ranganath, Ricardo Henao, and Russell Greiner. An effective meaningful way to evaluate survival models. In *Proceedings of the 40th International Conference on Machine Learning (ICML)*, volume 202 of *PMLR*, pp. 28244–28276, 2023a.

Shi-ang Qi, Neeraj Kumar, Ruchika Verma, Jian-Yi Xu, Grace Shen-Tu, and Russell Greiner. Using Bayesian neural networks to select features and compute credible intervals for personalized survival prediction. *IEEE Transactions on Biomedical Engineering*, 70(12):3389–3400, 2023b. doi: 10.1109/TBME.2023.3287514.

Shi-ang Qi, Yakun Yu, and Russell Greiner. Conformalized survival distributions: a generic post-process to increase calibration. In *Proceedings of the 41st International Conference on Machine Learning (ICML)*, volume 235 of *PMLR*, 2024.

Pengzhen Ren, Yun Xiao, Xiaojun Chang, Po-Yao Huang, Zhihui Li, Brij B. Gupta, Xiaojiang Chen, and Xin Wang. A survey of deep active learning. *ACM Computing Surveys*, 54(9):1–40, 2021.

David Rindt, Robert Hu, David Steinsaltz, and Dino Sejdinovic. Survival regression with proper scoring rules and monotonic neural networks. In *Proceedings of the 25th International Conference on Artificial Intelligence and Statistics (AISTATS)*, volume 151 of *PMLR*, 2022.

Maxim Sviridenko. A note on maximizing a submodular set function subject to a knapsack constraint. *Operations Research Letters*, 32(1):41–43, 2004.

Matt Taddy, Chun-Sheng Chen, Jun Yu, and Mitch Wyle. Bayesian and empirical Bayesian forests. In *Proceedings of the 32nd International Conference on Machine Learning*, 2015.

Anthony Joe Turkson, Francis Ayiah-Mensah, and Vivian Nimoh. Handling censoring and censored data in survival analysis: a standalone systematic literature review. *International Journal of Mathematics and Mathematical Sciences*, 2021:9307475, 2021. doi: 10.1155/2021/9307475.

Bhanukiran Vinzamuri, Yan Li, and Chandan K. Reddy. Active learning based survival regression for censored data. In *Proceedings of the 23rd ACM International Conference on Information and Knowledge Management (CIKM)*, pp. 241–250, 2014. doi: 10.1145/2661829.2662065.

Chun-Nam Yu, Russell Greiner, Hsiu-Chin Lin, and Vickie Baracos. Learning patient-specific cancer survival distributions as a sequence of dependent regressors. In *Advances in Neural Information Processing Systems 24 (NeurIPS)*, 2011.

# A Theoretical Analysis

## A.1 Notation

Table 3: Symbols used throughout the paper.

| Symbol | Meaning | Defined in |
|---|---|---|
| *Data and dataset* | | |
| $x_i$ | Covariate vector of instance $i$ | Sec. 3 |
| $t_i$ | Recorded time of instance $i$ (event time if uncensored, lower bound if censored) | Sec. 3 |
| $\delta_i$ | Event indicator (1 event observed, 0 censored) | Sec. 3 |
| $c_i$ | Probing cost of instance $i$ | Sec. 3 |
| $D$ | Training dataset $\{[x_i, t_i, \delta_i, c_i]\}_{i=1}^{L}$ | Sec. 3 |
| $L$ | Number of training instances | Sec. 3 |
| $B$ | Total budget | Sec. 3 |
| $k$ | Probe depth (years of additional follow-up the oracle returns) | Sec. 3 |
| $F$ | Set of indices selected for probing | Sec. 3 |
| $T_i$ | Random variable for the true (unobserved) event time of instance $i$ | Sec. 4.2 |
| $C$ | Censoring time in $T \perp C \mid X$; distinct from the objective $C(S)$ | Sec. 3 |
| $t_i^{\text{probed}}$, $\delta_i^{\text{probed}}$ | Oracle output for instance $i$ after probing | Sec. 5 |
| *Time discretisation* | | |
| $n$ | Number of time bins | Sec. 3 |
| $b_j$ | The $j$-th bin | Sec. 3 |
| $y$ | Bin-valued classification label of an instance ($y \in \{b_1, \ldots, b_n\}$) | Sec. 3 |
| $j_i$ | Bin index of $t_i$ (the bin containing the censoring time of instance $i$) | Sec. 4.2 |
| *Model and Bayesian inference* | | |
| $\omega$ | Model parameters | Sec. 4 |
| $M$ | Bayesian model | Sec. 3 |
| $q_\phi(\omega)$ | Variational posterior over $\omega$ with parameters $\phi$ | Sec. 4 |
| $\hat{\omega}_k$ | $k$-th posterior sample, $k = 1, \ldots, K$ ($K$=64: four members $\times$ sixteen draws) | Sec. 4 |
| $\mathcal{L}(\phi)$ | ELBO objective | Sec. 4 |
| $p_{\text{cens}}(y = b_j \mid \omega)$ | Censoring-renormalised bin probability (Equation (1)) | Sec. 4.1 |
| *Acquisition functions and information theory* | | |
| $I(\cdot\,;\cdot \mid \cdot)$ | Conditional mutual information | Sec. 4 |
| $H(\cdot)$ | Entropy | Sec. 4 |
| $a_{\text{BBald}}$ | BatchBALD acquisition function (Equation (12)) | Sec. 4 |
| $\mathcal{B}$ | A batch (subset of instances under consideration) | Sec. 4.2 |
| $X_V, Y_v$ | Target covariate sample ($\lvert V\rvert = 300$) and the hypothetical label at $x_v$ | Sec. 4.3 |
| $f_a, \hat{c}_a$ | Population-targeted response information of action $a$ (Equation (4)) and its expected realised cost | Sec. 4.3 |
| $F_V(S), C(S)$ | Joint population-targeted objective and its entropy-capped surrogate | Sec. 4.3, App. A.7 |
| $Z_i$ | Finite-window event, natural-censoring or window-end response | Sec. 4.2 |
| *Combinatorial / coverage* | | |
| $\mathcal{U}$ | Ground set of elements (in coverage formulation) | App. A.4 |
| $\mathcal{S}$ | Family of subsets (in coverage formulation) | App. A.4 |
| $w(\cdot)$ | Element weight (in weighted maximum coverage) | App. A.4 |
| $D_{\text{pool}}$ | Pool of candidate instances in BLCP | App. A.5 |
| $\mathcal{I}$ | Submodular information oracle in BLCP | App. A.5 |

## A.2 Bayesian pipeline details

With this adjusted distribution, standard information-theoretic acquisition functions apply directly. In particular, the BatchBALD algorithm (Kirsch et al., 2019) is a state-of-the-art batch active learning method that selects an informative batch of $b$ instances by maximising the mutual information between their labels and the model parameters. Kirsch et al. (2019) show this MI decomposes as

$$I(y_{1:b}; \omega \mid x_{1:b}, D) = H(y_{1:b} \mid x_{1:b}, D) - \mathbb{E}_{p(\omega \mid D, x_{1:b})}\big[H(y_{1:b} \mid x_{1:b}, \omega, D)\big], \tag{9}$$

i.e. the predictive entropy minus the expected conditional entropy (we drop the conditioning on $x_{1:b}$ and $D$ for brevity). Both are estimated by Monte Carlo with $K$ posterior samples $\hat{\omega}_1, \ldots, \hat{\omega}_K$:

$$H(y_{1:b}) \approx -\sum_{\hat{y}_{1:b}} \bar{p}(\hat{y}_{1:b}) \log \bar{p}(\hat{y}_{1:b}), \quad \mathbb{E}_{p(\omega)}\big[H(y_{1:b} \mid \omega)\big] \approx \tfrac{1}{K} \sum_{i=1}^{b} \sum_{k=1}^{K} H(y_i \mid \hat{\omega}_k), \tag{10}$$

where $\bar{p}(\hat{y}_{1:b}) := \frac{1}{K}\sum_k p(\hat{y}_{1:b} \mid \hat{\omega}_k)$; the second estimator factorizes across the batch under the conditional-independence assumption BatchBALD already requires (full derivation below). The two terms cancel when posterior uncertainty is low (high confidence, low information gain) and diverge when it is high. BatchBALD was originally designed for classification with mutually exclusive classes and adapts to time-to-event tasks via the bin discretisation of Section 3.

This subsection expands the compact main-text description of our Bayesian downstream learner. We describe (a) the discrete-time MTLR likelihood under right censoring, (b) the prior alternatives we considered and why we default to spike-and-slab, (c) the variational ELBO objective used for posterior inference, and (d) the Monte Carlo procedure that turns posterior samples into the mutual-information estimates of Equations (9) and (10).

**Why a Bayesian downstream learner.** BatchBALD requires a measure of the uncertainty of the model's prediction for each candidate instance in order to compute mutual information (Kirsch et al., 2019). Bayesian models meet this requirement directly by maintaining a probability distribution over each parameter and propagating that distribution to the predictive output. The trade-off is computational cost (the parameter distributions must be fit and sampled) and a strong dependence on prior choice (Qi et al., 2023b); we discuss both below.

**Choice of survival framework.** Qi et al. (2023b) adapts and tests two Bayesian survival frameworks: **Cox Proportional Hazards**, which parameterises the hazard with a baseline plus covariate-dependent shift (Cox, 1972), and **Multi-Task Logistic Regression (MTLR)**, which discretises time into bins and fits a series of logistic regressions across them (Yu et al., 2011). We use the Bayesian MTLR variant of Qi et al. (2023b) because it (i) directly utilises censored observations during model training without imputation, (ii) requires no proportional-hazards assumption, (iii) provides per-instance event probability estimates within discrete time intervals (the form BatchBALD's MI computation expects), and (iv) maintains reasonable computational complexity despite its Bayesian formulation.

**Likelihood (Bayesian MTLR).** For instance $i$ with covariates $x_i$, the MTLR head outputs a vector of bin-level logits $\eta_i(\omega) \in \mathbb{R}^n$ that, after the standard MTLR construction, induces a categorical distribution over the time bins $\{b_1, \ldots, b_n\}$. An uncensored instance contributes $\log p(y_i = b_{j_i} \mid x_i, \omega)$ where $j_i$ is the event bin; a censored instance at bin $j_i$ contributes $\log \sum_{r=j_i}^{n} p(y_i = b_r \mid x_i, \omega)$, so right-censored points enter the likelihood without imputation.

**Prior alternatives.** The MTLR weights $\omega$ are equipped with a sparsity-inducing prior $p(\omega)$. Qi et al. (2023b) considers two:

- *Spike-and-slab*: a mixture of a concentrated spike and a diffuse slab encourages shrinkage of weakly supported weights. A point-mass spike permits exact zeros; a continuous spike concentrates weights near zero. These are distinct specifications, and the name alone does not imply exact feature exclusion.
- *Horseshoe*: $\omega_d \mid \tau, \lambda_d \sim \mathcal{N}(0, \tau^2\lambda_d^2)$ with half-Cauchy hyperpriors, which strongly shrinks small coefficients while leaving large ones essentially unpenalised.

We default to spike-and-slab in our experiments because the priors recommended in Qi et al. (2023b) were tuned for clinical survival data of the kind we use here.

**Posterior inference (ELBO).** Exact posterior inference is intractable, so we fit a variational posterior $q_\phi(\omega)$ by maximising the evidence lower bound:

$$\mathcal{L}(\phi) \;=\; \mathbb{E}_{q_\phi(\omega)}[\log p(\mathcal{D} \mid \omega)] - \mathrm{KL}(q_\phi(\omega) \,\|\, p(\omega))\,, \tag{11}$$

which trades data fit (the MTLR likelihood above) against deviation from the prior. After training, the model is summarised by $q_\phi(\omega)$.

**BatchBALD acquisition and MI-estimator details.** The BatchBALD acquisition function is (Kirsch et al., 2019)

$$a_{\mathrm{BBald}}(\{x_{1:b}\},\, p(\omega \mid \mathcal{D}_t)) = I(y_{1:b};\, \omega \,|\, x_{1:b}, \mathcal{D}_t)\,, \tag{12}$$

the conditional mutual information of the batch labels with the model parameters; it decomposes via Equation (9) and is estimated by the two Monte Carlo expressions of Equation (10), with $K{=}64$ posterior samples (four bootstrap members $\times$ sixteen draws) $\hat{\omega}_1, \ldots, \hat{\omega}_K$ drawn equally from the four fitted variational posteriors. Their equal-weight mixture is $\bar{q}(\omega) = \frac{1}{4}\sum_{m=1}^{4} q_{\phi_m}(\omega)$. The factorisation in the second estimator requires mutual conditional independence, $p(y_{1:b} \mid \omega, D) = \prod_i p(y_i \mid \omega, D)$; the first estimator does not factorise and is computed jointly over the batch. We apply the censoring renormalisation of Equation (1) to all bin probabilities $p(y_i = b_j \mid \hat{\omega}_k)$ *before* computing either term, giving the censoring-aware MI estimator.

### A.3 Positioning of this work

This work sits at the intersection of *survival analysis*, *active learning*, *budgeted learning*, and the *non-uniform-cost* variant of the latter, with a fifth axis (*probe depth*) that we introduce as a refinement of the partial-information aspect of survival data. We are not claiming that our method requires *all* of these axes simultaneously; rather, the formulation specialises gracefully when any subset is irrelevant (e.g. collapses to standard EPIG (Bickford Smith et al., 2023) on the event-time label when probe depth is infinite and costs are uniform, and to standard active learning when there is no budget). Table 4 positions this work against the closest published predecessors, and Figure 2 places it among the areas it draws on.

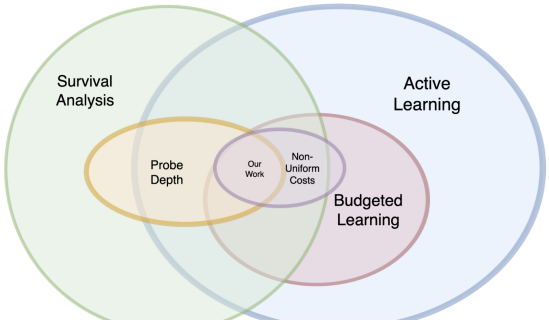



Figure 2: Positioning of the problem: probe depth specifies the follow-up acquired from survival data, while non-uniform costs determine how the acquisition budget is allocated.

Table 4: Comparison of relevant works across study aspects.

| Papers | Active Learning | Budget Constraints | Survival Analysis | Probe Depth | Non-uniform Costs |
|---|---|---|---|---|---|
| Lizotte et al. (2012) | ✓ | ✓ | – | – | – |
| Dedja et al. (2023) | ✓ | – | ✓ | partly (random) | – |
| Hüttel et al. (2025) | ✓ | – | ✓ | – | – |
| Bickford Smith et al. (2023) | ✓ | – | – | – | – |
| **Our Work** | ✓ | ✓ | ✓ | ✓ | ✓ |

### A.4 Maximum Coverage Problem

The *maximum coverage problem* is a well-known combinatorial problem that involves choosing $k$ sets from a group of $N$ sets of elements. The task is to choose the $k$ sets whose union has the maximum possible size. For example, if the sets are as follows: $S_1 = \{1, 2, 3\}$, $S_2 = \{2, 3, 4\}$, $S_3 = \{4, 5\}$, $S_4 = \{6\}$, where here $N = 4$ and $k = 2$, then the optimal choice of sets here are sets $S_1$ and $S_3$, as their union $\{1, 2, 3, 4, 5\}$ is larger than the union of $S_1$ and $S_2$: $\{1, 2, 3, 4\}$ or $S_2$ and $S_3 = \{2, 3, 4, 5\}$, and all pairs with $S_4$ make at most 4. This problem is known to be **NP**-hard to solve optimally.

There is an extension of this problem called the *weighted maximum coverage problem* where everything is the same except each integer has attached to it a weight, and the goal is now to maximize the sum of the weight of the union rather than simply the size of the union. In the example above, if the integers 1,2,3,4, and 5 all had weight 1, but 6 had weight 10, then now we certainly would wish to

include $S_4$ as part of one of the sets we choose, in this case you could choose $S_1$ and $S_4$, or $S_2$ and $S_4$, as both would give you a total highest weight of 13.

Figure 3 illustrates coverage as a union of weighted regions. This is intuition for redundant information, not a claim that every mutual-information objective has a set-union representation. Our formal reduction runs from weighted maximum coverage to the entropy-capped objective (Theorem A.4); submodularity supplies the approximation guarantee separately.

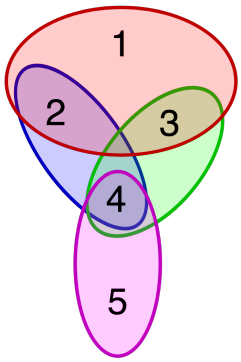



Figure 3: Coverage intuition: selecting shapes covers the union of their weighted regions. The diagram is not a general representation theorem for mutual information.

**Definition A.1.** A set function $f : 2^N \to \mathbb{R}$ defined on the subsets of a finite set $N$ is called **submodular** if for every $A, B \subseteq N$,

$$f(A) + f(B) \geq f(A \cup B) + f(A \cap B)$$

Equivalently, $f$ is submodular if it satisfies the **diminishing returns** property: for every $A \subseteq B \subseteq N$ and $x \in N \setminus B$,

$$f(A \cup \{x\}) - f(A) \geq f(B \cup \{x\}) - f(B)$$

A normalised submodular function satisfies $f(A) + f(B) \geq f(A \cup B)$ for disjoint $A$ and $B$. Parameter information has this property under mutual conditional independence of responses given the parameters; prediction-targeted information need not. The capped surrogate $C$ is submodular because each term is a concave, non-decreasing function of a non-negative modular sum, independently of whether candidate responses are conditionally independent.

The greedy approach for the maximum coverage problem involves iteratively selecting the element or set that provides the largest marginal gain in coverage at each step. It gives the highest known guaranteed lower bound of all polynomial-time approximation algorithms. In fact, a polynomial-time algorithm with a better constant for maximum coverage would imply **P = NP** (Feige, 1998). The greedy algorithm discussed above achieves a lower bound of $1 - 1/e$ (around 63%) of the optimal, and is the algorithm used in BatchBALD for selecting its batch (Kirsch et al., 2019).

### A.5 Reductions of the Computational Problem

#### A.5.1 Reducing the Weighted Maximum Coverage Problem to the Budgeted Learning Computation Problem

In this section, we formally reduce the weighted maximum coverage problem (WMCP), under the same initial assumptions outlined in the problem formulation, to the Budgeted Learning Computation Problem (BLCP). The difference between the Budgeted Learning problem and BLCP is that BLCP assumes the existence of an oracle $I$ with submodular and efficient computation properties that can assign an information value to any batch of data instances $B$. In the context of the paper, this role

is played by the entropy-capped objective $C$ of $\mathrm{EPIG}_{\mathrm{surv}}$ (Proposition A.3). Assuming we have this oracle, selecting the future batch of instances to learn about becomes a purely computational problem.

**The Budgeted Learning Computation Problem is defined as:**

Let $D_{pool} = \{x_1, \ldots, x_n\}$ be a pool of $n$ instances, each endowed with a cost $c_i > 0$, and let $B > 0$ denote the total budget. Assume the existence of a polynomial-time oracle that evaluates a normalised, non-negative, monotone submodular function

$$\mathcal{I} : 2^{D_{\text{pool}}} \longrightarrow \mathbb{R}_{\geq 0}.$$

Our objective is to select a subset

$$Q \subseteq D_{pool}$$

that maximizes $\mathcal{I}(Q)$ while respecting the knapsack constraint

$$\sum_{x_i \in Q} c_i \ \leq \ B$$

**The Weighted Maximum Coverage problem is defined as:**

Let $\mathcal{U} = \{e_1, \ldots, e_m\}$ be a ground set of elements. Each element $e_i$ carries a non-negative weight $w(e_i) \in \mathbb{R}_{\geq 0}$. Consider a family of subsets $\mathcal{S} = \{S_1, \ldots, S_n\} \subseteq 2^{\mathcal{U}}$. We can assume here that $|U| = poly(|\mathcal{S}|)$.

The weighted maximum coverage problem seeks a set

$$\mathcal{A} \subseteq \mathcal{S}$$

where $|\mathcal{A}| = K$ for some $K \in \mathbb{Z}_+$, that maximises the total weight of the elements it covers,

$$\max_{\mathcal{A}} \big\{ W(\mathcal{A}) \ = \sum_{e \in \cup_{S \in \mathcal{A}} S} w(e) \big\}.$$

**Theorem A.2.** *The weighted maximum coverage problem can be reduced in polynomial time to the Budgeted Learning Computation Problem assuming the cost for all instances is the same.*

*Proof.* Since we assume the cost is the same, the BLCP problem now involves choosing $K = \lfloor B/c \rfloor$ instances. Now for the reduction, for a given instance of inputs to the WMCP: $\mathcal{U}, \mathcal{S}$, we have to find a corresponding set of instances $D_{pool}$ and function $\mathcal{I}$ such that solving this BLCP means solving the WMCP instance. This can be done by:

1. $D_{pool} = \{x_1, \ldots, x_{|\mathcal{S}|}\}$
2. For any subset $Q \subseteq D_{pool}$, $\mathcal{I}(Q) = W(\mathcal{A})$ where $\mathcal{A} = \{S_i \mid x_i \in Q\}$

In the first point we ensure that the number of instances in BLCP equals the number of sets in WMCP. In the second point we can set our information function $\mathcal{I}$ to be equal to the weight of the union of elements of corresponding sets in WMCP. Notice this does not break any of our assumptions about $\mathcal{I}$ as $W$ is both submodular and computed in polynomial time.

□

The reduction identifies weighted coverage as a special case of the optimisation class. It does not establish that every information objective is a coverage function, or that every non-negative information matrix is attainable by a survival model. Survival response modelling determines the acquisition values; the cardinality or cost budget determines the feasible subsets.

**Algorithm 1** Marginal-density heuristic on $C$ for a cost budget; distinct from evaluated singleton ranking

**Input:** Budget $B$, pool $\mathcal{D}$, costs $\{c_i\}$, posterior $p(\omega \mid \mathcal{D}_{\text{train}})$
$A \leftarrow \emptyset$, cost $\leftarrow 0$
$D \leftarrow \{d_i \in \mathcal{D} \mid c_i \leq B - \text{cost}\}$
**while** $D \neq \emptyset$ **do**
  $d_j \leftarrow \arg\max_{d_i \in D} [C(A \cup \{d_i\}) - C(A)]/c_i$
  $A \leftarrow A \cup \{d_j\}$
  cost $\leftarrow$ cost $+ c_j$
  $D \leftarrow D \setminus \{d_j\}$
  $D \leftarrow \{d_i \in D \mid c_i \leq B - \text{cost}\}$
**end while**
**Output:** $A$

**Algorithm 2** Partial-enumeration $(1 - 1/e)$ approximation for Budgeted Maximum Coverage

Pool of points $S$, budget $B$, weights $w(\cdot)$, costs $\{c_i\}$, subset size $z$
$H_1 \leftarrow \arg\max\{w(G) : G \subseteq S,\ |G| \leq z,\ c(G) \leq B\}$
$H_2 \leftarrow \emptyset$
**for all** $G \subseteq S$ such that $|G| = z$ and $c(G) \leq B$ **do**
  $U \leftarrow S \setminus G$
  **while** $U \neq \emptyset$ **do**
    $x^\star \leftarrow \arg\max_{x \in U} \frac{\Delta w(x|G)}{c_x}$
    **if** $c(G) + c_{x^\star} \leq B$ **then**
      $G \leftarrow G \cup \{x^\star\}$
    **end if**
    $U \leftarrow U \setminus \{x^\star\}$
  **end while**
  **if** $w(G) > w(H_2)$ **then**
    $H_2 \leftarrow G$
  **end if**
**end for**
**if** $w(H_1) > w(H_2)$ **then**
  **Output:** $H_1$
**else**
  **Output:** $H_2$
**end if**

### A.6 Changing the Budgeted Maximum Coverage Algorithm

Khuller et al. (1999) provide a novel greedy algorithm for the budgeted maximum coverage case that still provides the same lower bound guarantees as the original greedy algorithm. We illustrate this algorithm in Algorithm 2.

For budgeted maximum coverage, $z = 3$ suffices for the stated bound (Khuller et al., 1999); partial enumeration also gives $(1 - 1/e)$ for monotone submodular maximisation under a single knapsack constraint (Sviridenko, 2004). Algorithm 1 omits enumeration. The evaluated rule additionally replaces marginal gain by singleton gain and imposes a batch-size limit and one action per patient (Algorithm 3). We do not transfer the single-knapsack guarantee to these additional constraints or to uncertain costs.

### A.7 The entropy-capped objective

Write $A_{va} := I(Y_v; Z_a \mid x_v, D)$, $h_v := H(Y_v \mid x_v, D)$ and $e_v := I(Y_v; \omega \mid x_v, D) \leq h_v$. The modular sum is $M(S) = \sum_{a \in S} f_a = \sum_v \sum_{a \in S} A_{va}$; the capped objective is $C(S) = \sum_v \min\{\sum_{a \in S} A_{va}, h_v\}$ (Equation (7)).

**Algorithm 3** Evaluated batch selection at a fixed posterior

1: **Input:** remaining budget $B_r$, batch limit $b$, eligible actions $\mathcal{A}$, singleton scores $f_a$, expected costs $\hat{c}_a$
2: Set reserve $u_a = c_i$ for known costs; $u_a = c_0 + k$ for realised cost and $a = (i, k)$
3: $S \leftarrow \emptyset$, $r \leftarrow 0$
4: **while** $|S| < b$ and some $a \in \mathcal{A}$ has $u_a \leq B_r - r$ **do**
5: $\quad a^\star \leftarrow \arg\max_{a \in \mathcal{A}:\, u_a \leq B_r - r} f_a / \hat{c}_a$
6: $\quad S \leftarrow S \cup \{a^\star\}$; $r \leftarrow r + u_{a^\star}$
7: $\quad$ Remove all actions for the selected patient from $\mathcal{A}$
8: **end while**
9: Query $S$, charge realised costs, update eligible follow-up and refit the learner

**Why use a surrogate.** $F_V(S) = \sum_v I(Y_v; Z_S \mid x_v, D)$ is not submodular even under Equation (6): with two independent fair parameter bits $U, W$, probes $Z_a = U$ and $Z_b = W$, and one target $Y = U \oplus W$, each probe alone carries 0 bits about $Y$ while the pair carries 1 bit, so the gain of $b$ grows after $a$ is chosen (Krause & Guestrin, 2005). This example gives $C < F_V$; conversely, with one target $Y = (U, W)$ of two fair bits and two probes that both reveal $U$, $C = M = 2$ bits while $F_V = 1$ bit, so $C > F_V$. Neither $M$ nor $C$ therefore bounds $F_V$ from either side; the guarantees below are relative to $\max C$.

**Proposition A.3.** *$C$ is normalised, monotone and submodular; $C(\{a\}) = f_a$ for every action; and if $\sum_{a \in S} A_{va} \leq h_v$ for every target and every feasible batch (for which $b\, e_v \leq h_v$ suffices), greedy on $C$ is top-$b$ selection by $f_a$.*

*Proof. Submodularity.* $x \mapsto \min\{x, h_v\}$ is concave, non-decreasing and 0 at 0; $x_v(S) = \sum_{a \in S} A_{va}$ is modular and non-decreasing because $A_{va} \geq 0$. A concave non-decreasing function of such a set function is normalised, monotone and submodular, and a sum over $v$ preserves all three.

*Single probes.* $A_{va} \leq h_v$ since mutual information is at most the entropy of either argument, so $C(\{a\}) = \sum_v A_{va} = f_a$.

*Certificate.* If no cap binds on any feasible batch, every marginal gain of $C$ equals $f_a$, and greedy on a modular objective under $|S| \leq b$ picks the $b$ largest $f_a$. Because $Y_v$ and $Z_a$ are conditionally independent given $\omega$, data processing gives $A_{va} \leq e_v$, hence $\sum_{a \in S} A_{va} \leq b\, e_v$. □

Under a cardinality constraint with one candidate action per patient, greedy therefore attains $C(S_{\text{greedy}}) \geq (1 - (1 - 1/b)^b) \max_{|S| \leq b} C(S) \geq (1 - 1/e) \max_{|S| \leq b} C(S)$ (Nemhauser et al., 1978). The certificate is checkable each round: when every target's epistemic share $e_v / h_v$ is at most $1/b$, the capped and uncapped rules select the same batch. The reported runs use the top-$b$ rule (Section 5); a weaker, per-batch certificate suffices for them: if no target's cap binds on the selected top-$b$ batch $S_{\text{top}}$, then $C(S) \leq M(S) \leq M(S_{\text{top}}) = C(S_{\text{top}})$ for every feasible $S$, so $S_{\text{top}}$ maximises $C$ exactly. Capped greedy is the variant that carries the guarantee; top-$b$ is the evaluated implementation.

**Theorem A.4.** *Weighted maximum coverage (Appendix A.5.1) reduces in polynomial time to maximising $C$ under a cardinality constraint, over instances given by an arbitrary non-negative information matrix $(A_{va})$ and target entropies $(h_v)$ with $A_{va} \leq h_v$. Hence no polynomial-time algorithm guarantees a constant above $1 - 1/e$ for this class unless $\mathrm{P} = \mathrm{NP}$. The class is that of the objective; we do not show that every such instance is realised by a survival model.*

*Proof.* Given elements $e$ with weights $w(e)$, sets $S_1, \ldots, S_n$ and cardinality $K$, take one target $v_e$ per element with $h_{v_e} = w(e)$, one action $a_i$ per set, and $A_{v_e a_i} = w(e)$ if $e \in S_i$ and 0 otherwise.

Then $\min\{\sum_{a_i \in Q} A_{v_e a_i}, h_{v_e}\}$ equals $w(e)$ if some chosen set contains $e$ and 0 otherwise, so $C(Q)$ is exactly the covered weight; the construction is linear in the instance size.

Maximum coverage, and hence weighted maximum coverage, admits no polynomial-time constant above $1 - 1/e$ unless $\mathrm{P} = \mathrm{NP}$ (Feige, 1998), and the reduction carries the bound over. □

### A.8 Expected and realised cost of an action

For action $a = (i, k)$ on a patient unresolved at observed time $t_i$, the realised cost charged when the probe is answered is $c_0 + \min(T_i, C_i^{\mathrm{nat}}, t_i + k) - t_i \le c_0 + k$, where $C_i^{\mathrm{nat}}$ is the natural-censoring time; its expectation given the current history is $c_0 + \int_{t_i}^{t_i+k} P(T_i > u,\, C_i^{\mathrm{nat}} > u \mid T_i > t_i,\, C_i^{\mathrm{nat}} > t_i,\, x_i, D)\, du$. We estimate it by the product form $\hat{c}_a = c_0 + \int_{t_i}^{t_i+k} P(T_i > u \mid T_i > t_i, x_i, D)\, \hat{P}(C_i^{\mathrm{nat}} > u \mid C_i^{\mathrm{nat}} > t_i)\, du$, which additionally assumes that event and natural-censoring times are independent given the history. The first factor is the ratio of posterior-mean survival at $u$ and at $t_i$; the second is a pooled (covariate-free) reverse Kaplan–Meier estimate of the natural-censoring survival on the development set, in which a probe that closes without an endpoint is treated as right-censoring of the censoring process; both are evaluated bin-wise on the time grid. $\hat{c}_a$ is a plug-in estimate with no calibration guarantee. Selection reserves $c_0 + k$ per accepted action and checks feasibility against the reserved total; $\hat{c}_a$ is used only for ranking, so by induction over rounds the realised cumulative cost never exceeds $B$. With random costs, $\hat{c}_a$ is replaced by the instance's drawn cost $c_i$, which is known to the learner; with uniform costs it is 1.

## B Sequential guarantees and their limits

**Setting.** $\Omega$ is a finite hypothesis set with prior $p(\omega)$; every probe $a$ returns $Z_a \sim p(z \mid \omega)$, responses conditionally independent given $\omega$; the targets are $Y_v \sim p(y \mid x_v, \omega)$. A policy $\pi$ spends $K$ probes in $R$ rounds of $b = K/R$ with exact conditioning between rounds. Its value is $\mathcal{V}(\pi) = \sum_v I(Y_v; \mathrm{Tr}_\pi)$, the expected whole-run reduction of target log-loss, where $\mathrm{Tr}_\pi$ is the transcript of probes and responses; $\mathrm{OPT}_K$ is the supremum over adaptive policies. $\mathcal{C}_L$ is the class of policies whose choice at each round depends only on observed history, internal randomness and oracle access to current joint laws of $(Y_V, Z_S)$ for $|S| \le L$, without access to the latent model representation or hidden placement of informative probes: $\mathrm{EPIG}_{\mathrm{surv}}$ is in $\mathcal{C}_1$, its batch and capped forms in $\mathcal{C}_b$, any $L$-step lookahead in $\mathcal{C}_L$. At history $h$ write $d(a \mid h) = \sum_v I(Y_v; Z_a \mid h)$, likewise $d(S \mid h)$, and $m_h = \max_a d(a \mid h)$.

**Theorem B.1** (No unconditional guarantee for bounded-lookahead target-focused rules)**.** *For every $L \ge 1$, $b$, $\varepsilon > 0$, $K \ge L + 1$ and $\pi \in \mathcal{C}_L$ there is an instance with $\mathcal{V}(\pi) \le \varepsilon\, \mathrm{OPT}_K$.*

*Proof. Instance.* $\omega$ consists of $m = L+1$ special fair bits $U_i$ and $N$ decoy fair bits $V_j$, all independent; the single target is $Y = U_1 \oplus \cdots \oplus U_m$; probe $i$ of the $m$ specials reveals $U_i$, probe $j$ of the $N$ decoys reveals $V_j$; the specials are placed uniformly at random among the $n = m + N$ candidates. $\mathrm{OPT}_K = 1$ bit, by buying the specials.

*Blindness.* Any set of fewer than $m$ specials leaves $Y$ exactly fair, so before the first special is bought every joint law of $(Y, Z_S)$ with $|S| \le L$ is the same for every $S$, and the policy's picks are independent of the placement.

*Bound.* $\mathcal{V}(\pi) \le P(\text{all } m \text{ specials bought}) \le Km/n$, which is at most $\varepsilon$ for $n \ge Km/\varepsilon$. The bound holds for every batch size $b$. Averaging over placements therefore implies that at least one fixed placement satisfies the stated inequality. □

The same instance also defeats rules based solely on parameter-information scores: every candidate carries exactly one fair bit of parameter information. A rule with access to the model's structure identifies the specials, and an exhaustive planner over response sequences attains $\mathrm{OPT}_K$, at a cost exponential in $K$.

**Theorem B.2** (Whole-run guarantee under bounded growth)**.** *Assume **(A1)**: for every reachable history $h$, every extension $h'$ by at most $K$ observations and every unobserved probe $a$, $d(a \mid h') \le \kappa\, m_h$. Then greedy $\mathrm{EPIG}_{\mathrm{surv}}$ (top-$b$ on $d(\cdot \mid h)$, or its capped form) satisfies $\mathcal{V} \ge \big[1 - (1 - \frac{1}{\kappa K})^R\big] \mathrm{OPT}_K$.*

*Proof. Retention.* The batch chosen at $h_t$ contains $a^\star = \arg\max_a d(a \mid h_t)$ (for the capped form too, since $C(\{a\}) = f_a$), so by the chain rule its gain is at least $m_{h_t}$.

*Growth.* Append an optimal $K$-probe policy to the greedy history, simulating that policy's decisions from its own transcript and reusing any responses already observed. The combined transcript contains

the optimal policy's transcript, so its expected information is at least $\mathrm{OPT}_K$. By the chain rule and A1, the additional information in at most $K$ observations is at most $K\kappa m_{h_t}$ conditional on $h_t$. Averaging over greedy histories gives $\mathrm{OPT}_K - \mathcal{V}_t \le \kappa K\, \mathbb{E} m_{h_t}$.

*Recurrence.* Hence $\mathcal{V}_{t+1} - \mathcal{V}_t \ge (\mathrm{OPT}_K - \mathcal{V}_t)/(\kappa K)$; iterating $R$ times from $\mathcal{V}_0 = 0$ gives the bound. □

A1 is exactly what the instance of Theorem B.1 violates (a special probe's gain jumps from 0 to positive), so some condition excluding that instance is needed for any guarantee of this kind, and A1 is one sufficient condition. At the paper's scale ($K = 400$, $b = 10$, $R = 40$) and $\kappa = 1$ the factor is $0.095$; we have not certified $\kappa$ on the fitted models. Every statement in this appendix was also checked by exact enumeration on small instances.

**Convergence.** The acquisition rule decides which patients are followed up, not how the model is fitted: if every policy exhausts each pool patient's available follow-up, the resulting observed datasets are identical. Identical fitted models additionally require a fitting procedure whose output does not depend on the acquisition path; convergence alone does not ensure this. Our warm-started refits may retain path dependence. Theorem B.2 bounds information acquisition under bounded growth; it does not establish statistical consistency of the learner.

**Misspecification.** $\mathrm{EPIG}_{\mathrm{surv}}$ maximises information about the model's own predictions, so its guarantees are relative to the model. The real-data experiments do not establish that linear MTLR is correctly specified. They provide empirical evidence under possible misspecification, complemented by results with a structurally different learner (Appendix H).

## C Synthetic studies with an exact posterior

The real-data results reflect both acquisition decisions and their interaction with approximate model fitting. These studies isolate acquisition using a finite hypothesis space containing the data-generating model and exact Bayesian posterior updates. Performance is evaluated by the exact KL divergence from the true target law to the posterior predictive distribution, without test-sample or Monte Carlo error. Each cell below is the mean over seeds of the area under that KL curve divided by its horizon ($\times 100$; lower is better), the primary endpoint of a pre-registration written before any rule was compared; every rule receives the same posterior and the same response probabilities.

**Seven pre-registered worlds** (40 seeds each; Table 5): *matched* is the uniform-cost setting (one depth, equal costs, targets drawn like the pool); *nuisance* adds two parameter directions with a wide prior that the targets do not depend on; *heavy cens.* raises the natural-censoring hazard so that most probes return no endpoint; *shift* moves the target covariates away from the pool; *saturated* adds 400 fully observed records, leaving little to learn; *random prices* and *realised cost* are the random-cost and realised-follow-up-cost settings, with every rule divided by its (expected) cost. Paired over seeds, $\mathrm{EPIG}_{\mathrm{surv}}$ has no resolved loss in these worlds (22W/27T/0L over 49 comparisons). Where a baseline's mean is lower (matched, heavy cens., shift, realised cost) the difference is unresolved; in the saturated world, all comparisons are unresolved, consistent with the limited scope for improvement. The advantage is largest early: in the matched world $\mathrm{EPIG}_{\mathrm{surv}}$ is ahead of Random after 20 probes but behind it at the final budget, illustrating that a one-step objective need not optimise final-budget performance.

**A world built from the assumptions of Section 4.3** (*showcase* column). The covariate is a cell index and each cell has its own hazard with its own independent posterior, so a probe informs its own cell and nothing else. Three cells hold the deployment population; four are *epistemic decoys* (very diffuse posterior, no targets), three are *aleatoric decoys* (known, mid-range hazard) and two are *dead* (known hazard near zero), each aimed at a different family of rules. Here $\mathrm{EPIG}_{\mathrm{surv}}$ beats every baseline (7W/0T/0L, 40 seeds) and spends 53% of its probes in the target cells.

**Batch size.** In the showcase world at an equal budget of about 200 probes, $\mathrm{EPIG}_{\mathrm{surv}}$ has the lowest value at batch sizes 1, 5, 10 and 20 (Table 6) and beats all 7 baselines at each (7W/0T/0L, 7W/0T/0L, 7W/0T/0L, 7W/0T/0L); its value rises slowly with the batch size, as the posterior is updated less often. All real-data runs use $b = 10$.

Table 5: Synthetic worlds with an exact posterior: area under the target-KL curve ($\times 100$, lower is better), mean $\pm$ standard error over seeds; best value per column in bold. Paired verdicts are given in the text.

| | matched | nuisance | heavy cens. | shift | saturated | random prices | realised cost | showcase |
|---|---|---|---|---|---|---|---|---|
| $\mathrm{EPIG}_{\mathrm{surv}}$ (ours) | 5.627 ±0.246 | **3.223** ±0.339 | 6.027 ±0.217 | 9.843 ±0.842 | **0.394** ±0.043 | **5.761** ±0.240 | 5.657 ±0.244 | **6.850** ±0.294 |
| C-BALD | 5.676 ±0.249 | 3.303 ±0.352 | 6.025 ±0.217 | 9.775 ±0.840 | 0.396 ±0.043 | 5.867 ±0.243 | 5.685 ±0.245 | 16.645 ±0.771 |
| Entropy | 5.763 ±0.243 | 3.520 ±0.340 | 6.038 ±0.216 | 10.048 ±0.894 | 0.401 ±0.044 | 6.119 ±0.234 | 5.716 ±0.235 | 16.607 ±0.728 |
| Variance | 5.967 ±0.246 | 3.512 ±0.366 | 6.153 ±0.222 | 10.478 ±0.915 | 0.405 ±0.043 | 6.203 ±0.241 | 6.012 ±0.246 | 14.764 ±0.689 |
| IDEAL | 5.967 ±0.246 | 3.512 ±0.366 | 6.153 ±0.222 | 10.478 ±0.915 | 0.405 ±0.043 | 6.187 ±0.241 | 6.014 ±0.246 | 13.163 ±0.576 |
| Dedja | 5.642 ±0.239 | 3.369 ±0.337 | **6.003** ±0.212 | 9.840 ±0.852 | 0.395 ±0.042 | 6.153 ±0.231 | 5.823 ±0.241 | 16.435 ±0.722 |
| BatchBALD | 6.788 ±0.228 | 4.107 ±0.296 | 6.780 ±0.210 | 11.874 ±0.869 | 0.418 ±0.045 | 7.218 ±0.242 | 6.101 ±0.232 | 15.386 ±0.726 |
| Random | **5.583** ±0.225 | 3.468 ±0.309 | 6.098 ±0.205 | **9.673** ±0.768 | 0.401 ±0.043 | 6.270 ±0.226 | **5.635** ±0.218 | 11.237 ±0.491 |

Table 6: Showcase world, batch sizes 1–20 at an equal budget of about 200 probes, 40 seeds: area under the target-KL curve ($\times 100$, lower is better), mean $\pm$ standard error over seeds; best value per column in bold.

| | $b = 1$ | $b = 5$ | $b = 10$ | $b = 20$ |
|---|---|---|---|---|
| $\mathrm{EPIG}_{\mathrm{surv}}$ (ours) | **6.080**±0.259 | **6.244**±0.271 | **6.850**±0.294 | **8.064**±0.352 |
| C-BALD | 14.618±0.676 | 15.533±0.693 | 16.645±0.771 | 18.046±0.846 |
| Entropy | 16.282±0.711 | 16.480±0.718 | 16.607±0.728 | 16.992±0.747 |
| Variance | 12.022±0.563 | 13.204±0.618 | 14.764±0.689 | 15.746±0.781 |
| IDEAL | 11.185±0.485 | 11.513±0.494 | 13.163±0.576 | 14.493±0.688 |
| Dedja | 16.102±0.706 | 16.380±0.713 | 16.435±0.722 | 16.823±0.737 |
| BatchBALD | 14.966±0.728 | 14.639±0.705 | 15.386±0.726 | 16.347±0.805 |
| Random | 11.191±0.491 | 10.590±0.453 | 11.237±0.491 | 11.846±0.512 |

**Cost and depth.** Two further worlds add the ingredients of the later settings (40 seeds each; Table 7). With an independent price per patient and every rule divided by price, $\mathrm{EPIG}_{\mathrm{surv}}$ beats every baseline (7W/0T/0L). With a menu of depths $\{1, 3, 6\}$ and realised follow-up cost, where $\mathrm{EPIG}_{\mathrm{surv}}$ chooses the depth and each baseline probes at a fixed depth, the record is 4W/1T/2L: the depth-menu advantage of Appendix G does *not* reproduce here. Under a shared cost budget our rule buys deep windows and therefore far fewer probes than rules pinned to a shallow depth, and the cost scale was not calibrated to balance these alternatives. This result identifies a limitation of the density heuristic in this setting.

Table 7: Synthetic settings with heterogeneous costs: area under the target-KL curve against spend ($\times 100$, lower is better), mean $\pm$ standard error over seeds; best value per column in bold.

| | random prices | depth menu |
|---|---|---|
| $\mathrm{EPIG}_{\mathrm{surv}}$ (ours) | **9.833**±0.419 | 18.516±0.882 |
| C-BALD | 20.655±0.953 | 26.350±1.109 |
| Entropy | 12.848±0.570 | 11.444±0.663 |
| Variance | 17.816±0.864 | 25.054±1.054 |
| IDEAL | 16.818±0.787 | 23.398±0.992 |
| Dedja | 23.513±1.110 | **11.093**±0.633 |
| BatchBALD | 18.237±0.863 | 22.235±0.950 |
| Random | 22.437±0.986 | 17.948±0.730 |

**Agreement between capped and uncapped selection.** The same code base checks Proposition A.3 on generic, duplicated and tightly clustered candidates: the entropy cap changed the selected batch in 0 of 90 batches, so the capped and uncapped rules coincide there. These checks establish agreement only on the tested synthetic batches; they do not certify all real-data acquisition rounds.

## D BASELINE ALGORITHM DETAILS

**Why these baselines.** We selected a set of six comparators that span the natural axes a budgeted-AL method for survival data should be tested against: principled censored-AL acquisitions, principled regression-AL acquisitions, canonical uncertainty heuristics, and survival-specific selection rules. Concretely: **C-BALD** is a published BALD-style acquisition explicitly designed for right-censored data; **IDEAL** is a representative principled regression-AL acquisition that combines uncertainty with a diversity term; **Entropy Sampling** and **Variance Sampling** are canonical uncertainty baselines; **BatchBALD** is the standard principled batch acquisition; and **Dedja et al.** is the query strategy of the closest prior framework, built for survival labels that are disclosed incrementally. **Random** is included as an uninformed reference baseline. With non-uniform costs every method except Random ranks candidates by its score divided by the probe's cost (the expected cost under realised follow-up costs), exactly as $\mathrm{EPIG}_{\mathrm{surv}}$/cost does. The full definitions are below.

**Inverse Distance Weighting (IDEAL) Acquisition.** Inverse Distance Weighting extends traditional IDW (Bemporad, 2023) and computes a weighted combination of model uncertainty and feature-space exploration:

$$a(x) \;=\; s^2(x) + d \cdot z(x), \tag{13}$$

where $s^2(x)$ represents uncertainty through weighted variance of predictions, $z(x)$ is a diversity term based on distances to existing observations, and $d$ is a parameter controlling the exploration–exploitation trade-off. As for every baseline, with non-uniform costs we divide the IDEAL score by the probe's cost and select greedily by this ratio until the budget is used.

**Censoring-aware Bayesian Active Learning by Disagreement (C-BALD).** C-BALD (Hüttel et al., 2025) extends the BALD acquisition function to censored regression by computing the mutual information between the model parameters and the joint outcome consisting of the observed value and its censoring indicator. This yields an acquisition function that decomposes into information gained from the label conditional on censoring and information gained from observing the censoring status itself. As discussed in Section 5, C-BALD's joint MI is computed treating the oracle's $(t_i^{\mathrm{probed}}, \delta_i^{\mathrm{probed}})$ output as the "observed value and censoring indicator" pair its formulation expects. In our implementation, C-BALD scores singleton information about the finite-window response defined in Section 4.2; with equal costs, the batch contains the top-$b$ scores.

**Entropy Sampling.** Entropy sampling is a well-established technique in AL (Ren et al., 2021) that targets instances where the model's predictive distribution has the highest entropy:

$$H(y_i \mid x_i) \;=\; -\sum_{j=1}^{n} p_{\mathrm{cens}}(y_i = b_j \mid x_i) \log_2 p_{\mathrm{cens}}(y_i = b_j \mid x_i). \tag{14}$$

The acquisition function then takes the instances with the largest entropy values until the budget is used up. We use the censoring-renormalised $p_{\mathrm{cens}}$ from Section 4.1, so that probability mass already excluded by the current censor time does not contribute to predictive uncertainty.

**Variance Sampling.** Variance Sampling scores each instance by the variance, across the $K$ posterior draws, of its predicted mean event time under the censoring-renormalised distribution,

$$\mathrm{Var}(x) \;=\; \widehat{\mathrm{Var}}_k\Big[\textstyle\sum_{j=1}^{n} m_j \, p_{\mathrm{cens}}(y = b_j \mid x, \hat{\omega}_k)\Big], \tag{15}$$

where $m_j$ is the midpoint of bin $b_j$, and selects candidates in descending score order, dividing by expected cost when costs are non-uniform, subject to budget feasibility.

**Dedja et al.** Dedja et al. (2023) propose an active-learning framework for survival analysis in which the oracle discloses a queried instance's label incrementally: each query extends a censored instance's follow-up by a random fraction of its remaining time, and the instance stays in the pool, and can be queried again, until its label is fully disclosed. Their learner is a random survival forest modified to use such partially disclosed labels, and they study several query strategies. We use their uncertainty-based strategy, the survival analogue of margin sampling: as survival prediction has no decision boundary, it queries the instances whose prediction is closest to the population average of the training set,

$$x^\star \;=\; \arg\min_x \left|\mathcal{M}(x) - \bar{Y}_{\mathcal{T}}\right|, \tag{16}$$

where $\mathcal{M}(x)$ is the model's prediction for $x$ and $\bar{Y}_{\mathcal{T}}$ its average over the training set. In our harness $\mathcal{M}(x)$ is the posterior-mean predicted event time under the censoring-renormalised distribution of Section 4.1, and candidates are taken in order of closeness until the budget is used.

**BatchBALD.** BatchBALD (Kirsch et al., 2019) extends BALD (Houlsby et al., 2011) from single points to batches. BALD scores each candidate by the mutual information between its label and the model parameters, and taking the top $b$ of these scores can fill a batch with near-duplicates that carry the same information. BatchBALD instead scores the batch as a whole, by the joint mutual information $I(y_{1:b}; \omega \mid x_{1:b}, D)$ of Equation (12), which counts shared information once. It builds the batch greedily, adding at each step the candidate with the largest gain in joint mutual information; because this objective is monotone and submodular, under equal costs and conditionally independent responses, the greedy batch achieves at least a fraction $1 - 1/e$ of the optimal batch information (Kirsch et al., 2019; Nemhauser et al., 1978). In our harness the bin probabilities are censoring-renormalised (Section 4.1), the two terms of Equation (9) are estimated from the same $K{=}64$ posterior draws that every other rule uses (Equation (10)), and with non-uniform costs each greedy step ranks candidates by their gain per unit cost. BatchBALD measures information about the model parameters rather than about predictions on a target population, which is the distinction our rule is built on (Section 2).

**Random.** We include uniform random sampling from the candidate pool, subject to the same budget and cost constraints as the principled methods, as an uninformed reference baseline. Concretely, instances are drawn without replacement from $D_{\text{pool}}$ in uniformly random order; the first instance whose addition would exceed the remaining budget is skipped, and we continue until no instance fits.

## E PROTOCOL DETAILS

We explored three real-world survival datasets, with varying numbers of patients, numbers of features, and percentages of censored data: the *Study to Understand Prognoses Preferences Outcomes and Risks of Treatment* (SUPPORT; 9,105 patients, 42 features, 32% censored) (Knaus et al., 1995), where the goal is to predict time until death for seriously ill hospitalized adults; the *Assay of Serum Free Light Chain* study (FLCHAIN; 7,874 patients, 8 features, 72% censored) (Dispenzieri et al., 2012), where the goal is to predict all-cause mortality from a serum assay; and the *Medical Information Mart for Intensive Care IV* (MIMIC-IV; 38,319 patients, 91 features, 74% censored) (Johnson et al., 2023), a freely accessible, de-identified electronic health record database of patients admitted to the emergency department and intensive care units (ICUs) of the Beth Israel Deaconess Medical Center in Boston between 2008 and 2019, where the goal is to predict time until death from any cause after admission to the ICU. We use the all-cause-mortality survival cohort of MIMIC-IV introduced by Qi et al. (2023a): each patient contributes one ICU admission, and the 91 features are demographics and clinical measurements (vital signs and laboratory values) from the start of the stay. Dates of death come from both hospital and state records, so deaths after discharge are captured, and follow-up extends to one year after ICU admission.

**Detailed acquisition protocol.** Per seed we draw $n_{\text{dev}} = 3{,}000$ patients as the development set and up to 4,000 disjoint patients as the external test set. $n_0 = 160$ "historical" patients keep their recorded follow-up and provide the initial training and validation data ($n_{\text{val}} = 100$; 60 training rows with their recorded follow-up); every other development patient enters the training set artificially censored at an independent $C_i^{\text{art}} \sim U(0, q_{0.5})$, unless its recorded endpoint occurs first. This construction controls artificial censoring and creates scope for follow-up acquisition. Artificial censoring is independent of outcomes and covariates by construction. Together with the assumed independent

natural censoring and an acquisition policy based only on observed history, this supports the learner's censoring likelihood (Section 6.1); external NLL additionally relies on independent censoring in the test population. An earlier preprint's $t_i^{learner} = U_i t_i$ violated this independence ($\mathrm{Cov}(U_i T_i, T_i) = \frac{1}{2}\mathrm{Var}(T_i) > 0$). A uniform cut-off up to the median follow-up creates varied observation windows, approximating a registry with limited follow-up and leaving room for de-censoring to improve prediction; the level $q_{0.5}$ was fixed by a pre-declared headroom rule (Appendix E). The historical patients keep their full records, standing in for an established cohort whose outcomes are known when the new enrolment begins, and supply the initial training and validation data. A probe of patient $i$ to depth $k$ returns $\min(T_i, C_i^{\mathrm{nat}}, t_i + k)$ with its status; the patient remains eligible for further probes only while unresolved follow-up remains. The probe depth is the 25% quantile of residual survival among historical records, $k_{0.25}$. With uniform costs every method gets 400 probes; with random costs $c_i \sim U(0.2, 0.8)$ is drawn independently of $x_i$ and the outcome and the budget is 200 cost units; with realised follow-up costs a probe is charged $c_0 + \min(T_i, C_i^{\mathrm{nat}}, t_i + k) - t_i$ with $c_0 = k_{0.25}/2$ and budget $575\, k_{0.25}$; the variable-depth setting offers the depth menu $\{k_{0.1}, k_{0.25}, k_{0.5}\}$ under realised cost and budget $200\, k_{0.25}$ (Appendix Table 8). Time is discretised into $\sqrt{n_{\mathrm{events}}}$ bins; the learner is the Bayesian MTLR of Section 4.5; with uniform costs, $\mathrm{EPIG}_{\mathrm{surv}}$ selects the top-$b$ actions by $f_a$; with non-uniform costs, it ranks by $f_a/\hat{c}_a$ subject to budget feasibility. Top-$b$ maximises $C$ when no cap binds on the selected batch (Appendix A.7).

Table 8: The four settings. $c_0$ is overhead and $k$ is the requested depth; $T_i$ is the event time and $C_i^{\mathrm{nat}}$ the natural censoring time. The variable-depth confirmation uses $n_0 = 100$ historical patients with $n_{\mathrm{val}} = 40$ and artificial censoring up to $q_{0.9}$; the fixed-depth settings use $n_0 = 160$, $n_{\mathrm{val}} = 100$ and $q_{0.5}$. Every setting uses 60 training rows with their recorded follow-up, batches of $b = 10$ and a $4 \times 16$ ensemble.

| setting | depth | cost of probing $i$ | budget | seeds | where |
|---|---|---|---|---|---|
| uniform costs | one, $k_{0.25}$ | 1 | 400 probes | 40 dev + 40 fresh (2 datasets) | Section 6.1 |
| random costs | one, $k_{0.25}$ | $c_i \sim U(0.2, 0.8)$ | 200 units | 40 dev | Section 6.2 |
| realised follow-up costs | one, $k_{0.25}$ | $c_0 + \min(T_i, C_i^{\mathrm{nat}}, t_i + k) - t_i$ | $575\, k_{0.25}$ | 40 dev | Appendix F (extension) |
| variable depth | menu $\{k_{0.1}, k_{0.25}, k_{0.5}\}$ | realised, as above | $200\, k_{0.25}$ | 40 fresh | Appendix G (extension) |

**Censoring level.** The level $q_{0.5}$ of the fixed-depth settings was fixed by a rule declared before any acquisition run: the largest $q \in \{0.9, 0.7, 0.5\}$ at which every dataset has at least three paired SD of headroom (initial minus full-information NLL) and an initial C-index of at least $0.60$, with $q_{0.5}$ as the fallback. No level met it (FLCHAIN falls short at every level), so the fallback applied. The 40 development seeds also served an earlier robustness sweep, so only the fresh seeds are independent confirmations.

**Pre-registration of the variable-depth confirmation.** The candidate ($\mathrm{EPIG}_{\mathrm{surv}}$/cost, chosen on the 40-seed development pilots), the comparator roster (every baseline divided by expected cost, both at a fixed depth and over the depth menu), the primary endpoints (final external NLL and cost-indexed AULC; IBS and C-index secondary), the seed count (sized to resolve the weakest contrast of the development pilots, paired $d = 0.39$) and the pass rules were written down and frozen before the first confirmation seed was run on SUPPORT and FLCHAIN, and nothing was changed afterwards; MIMIC-IV was run later with every setting held fixed, although its own 40-seed pilot had favoured the fixed-depth variant of our method. The protocol's strictest gate — a positive mean paired difference against every comparator on final NLL and AULC — is not met in every cell: the depth-menu forms of some baselines on SUPPORT and MIMIC-IV have small negative means whose intervals contain zero, which the tables show as T.

## F Extension: realised follow-up costs at one depth

Charging each probe its realised follow-up instead of a random price (Table 9; paired AULC deltas in Figure 9) leaves the record intact: 7W/0T/0L on MIMIC-IV, 7W/0T/0L on SUPPORT and 7W/0T/0L on FLCHAIN against the seven baselines.

Table 9: Realised follow-up costs (development seeds, $n = 40$; budget $575\,k_{0.25}$ per seed, cost $c_0 + \min(T_i, C_i^{\text{nat}}, t_i + k) - t_i$ with $c_0 = 0.5\,k_{0.25}$; K1–K4 invariants pass on all three datasets). Mean paired AULC difference on the common fractional horizon (comparator − ours) with its 95% bootstrap interval, and verdict.

| comparator | MIMIC-IV | SUPPORT | FLCHAIN |
|---|---|---|---|
| C-BALD | $+0.0130$ [$+0.0053, +0.0211$] W | $+0.0028$ [$+0.0002, +0.0053$] W | $+0.0014$ [$+0.0008, +0.0021$] W |
| Entropy | $+0.1080$ [$+0.0819, +0.1351$] W | $+0.0460$ [$+0.0365, +0.0552$] W | $+0.0074$ [$+0.0049, +0.0098$] W |
| Variance | $+0.0804$ [$+0.0570, +0.1065$] W | $+0.0223$ [$+0.0163, +0.0284$] W | $+0.0059$ [$+0.0037, +0.0083$] W |
| IDEAL | $+0.0733$ [$+0.0575, +0.0897$] W | $+0.0267$ [$+0.0187, +0.0353$] W | $+0.0035$ [$+0.0014, +0.0056$] W |
| Dedja | $+0.0956$ [$+0.0738, +0.1158$] W | $+0.0248$ [$+0.0171, +0.0321$] W | $+0.0067$ [$+0.0044, +0.0092$] W |
| BatchBALD | $+0.1082$ [$+0.0809, +0.1349$] W | $+0.0429$ [$+0.0309, +0.0554$] W | $+0.0060$ [$+0.0039, +0.0082$] W |
| Random | $+0.0957$ [$+0.0718, +0.1187$] W | $+0.0190$ [$+0.0117, +0.0255$] W | $+0.0045$ [$+0.0021, +0.0069$] W |
| total (7 baselines) | 7W/0T/0L | 7W/0T/0L | 7W/0T/0L |

## G Extension: depth menu and realised costs (variable depth, confirmation)

The variable-depth setting offers the depth menu $\{k_{0.1}, k_{0.25}, k_{0.5}\}$ under realised follow-up cost with budget $200\,k_{0.25}$, on 40 fresh pre-registered seeds per dataset. It also uses $n_0 = 100$ historical patients and artificial censoring up to $q_{0.9}$ (Table 8), so it is a separate setting rather than an isolated depth ablation. The seven baselines probe at the middle menu depth and rank by their score divided by the probe's expected cost; $\text{EPIG}_{\text{surv}}$/cost chooses among the three depths and divides by expected cost. The frozen protocol also evaluated a depth-menu twin of each non-random baseline (score / expected cost over the whole menu), summarised below; Appendix E describes the pre-registration.

The full method — a menu of three depths, realised follow-up cost, information density — was confirmed on 40 fresh pre-registered seeds per dataset with candidate, roster and endpoints frozen beforehand (Table 10). Against the seven baselines $\text{EPIG}_{\text{surv}}$/cost has no resolved loss on NLL, AULC or IBS on any dataset: 20W/1T/0L, 19W/2T/0L, 17W/4T/0L, i.e. 56 wins, 7 unresolved, 0 losses over 63 comparisons; against the baselines' depth-menu twins, which share the depth menu and expected-cost model, none on NLL or AULC.

Table 10: Variable depth. Mean paired AULC difference (comparator − ours; positive favours ours) with its 95% bootstrap interval, and verdict.

| comparator | MIMIC-IV | SUPPORT | FLCHAIN |
|---|---|---|---|
| C-BALD | $+0.0084$ [$+0.0033, +0.0131$] W | $+0.0002$ [$-0.0002, +0.0006$] T | $-0.0001$ [$-0.0004, +0.0002$] T |
| Entropy | $+0.0473$ [$+0.0378, +0.0571$] W | $+0.0021$ [$+0.0012, +0.0032$] W | $+0.0023$ [$+0.0014, +0.0032$] W |
| Variance | $+0.0380$ [$+0.0292, +0.0470$] W | $+0.0021$ [$+0.0013, +0.0031$] W | $+0.0019$ [$+0.0012, +0.0026$] W |
| IDEAL | $+0.0372$ [$+0.0297, +0.0443$] W | $+0.0017$ [$+0.0010, +0.0025$] W | $+0.0015$ [$+0.0008, +0.0022$] W |
| Dedja | $+0.0428$ [$+0.0346, +0.0513$] W | $+0.0017$ [$+0.0008, +0.0027$] W | $+0.0021$ [$+0.0012, +0.0031$] W |
| BatchBALD | $+0.0481$ [$+0.0387, +0.0580$] W | $+0.0023$ [$+0.0014, +0.0032$] W | $+0.0022$ [$+0.0014, +0.0032$] W |
| Random | $+0.0485$ [$+0.0400, +0.0578$] W | $+0.0017$ [$+0.0009, +0.0026$] W | $+0.0024$ [$+0.0014, +0.0034$] W |
| total (7 baselines) | 7W/0T/0L | 6W/1T/0L | 6W/1T/0L |

**Comparisons with a shared depth menu.** Against the six depth-menu twins, $\text{EPIG}_{\text{surv}}$/cost is 17W/1T/0L on NLL and 16W/2T/0L on AULC.

## H Second learner: a random survival forest with Bayesian-bootstrap leaf hazards

**Learner.** Each ensemble member fits one random survival forest (100 trees, at least 20 patients per leaf, $\sqrt{d}$ candidate features per split) on the member's bootstrap rows, with times coarsened to the shared time grid: an event in bin $j$ is recorded at the bin's upper edge and a censoring at its lower edge, which is how the MTLR likelihood reads the bins. A posterior draw keeps the tree structure and re-estimates every leaf's discrete hazard on Dirichlet$(1, \ldots, 1)$-reweighted rows — one weight vector shared by all trees of the draw, so that the variation does not average away over trees (Taddy

et al., 2015) — with each leaf hazard shrunk toward the member-wide hazard by one pseudo-patient. Reweighting produces greater hazard variation in leaves with fewer patients at risk, providing the predictive uncertainty used by $\mathrm{EPIG}_{\mathrm{surv}}$. Four members with sixteen draws each give the same 64-draw tensor that every rule consumes with MTLR. The tree structure is refitted every five rounds and the leaf hazards every round; there is no warm start and the validation cohort is unused.

**How the learner was chosen.** The forest settings were selected by learner quality alone (initial and full-information NLL on development seeds) before any acquisition result, and frozen; the acquisition runs use fresh seeds. A first version with four-tree sub-forests as draws was discarded because its draws were tree noise rather than posterior uncertainty and its initial fit was far worse than MTLR's. A Bayesian Cox learner from the same framework (Qi et al., 2023b), linear or with one hidden layer, was tried first and is reported as a negative result: on SUPPORT and FLCHAIN even full outcome information did not improve its NLL, leaving little evidence of scope for acquisition to improve that fitted learner under this protocol.

**Results.** On SUPPORT and FLCHAIN the curves move (SUPPORT: $\mathrm{EPIG}_{\mathrm{surv}}$'s NLL 2.459→2.390; FLCHAIN 1.223→1.205) and Table 11 gives the final record; Table 12 shows it at every budget. The AULC ties on SUPPORT are C-BALD, Variance, BatchBALD, Random; on FLCHAIN, C-BALD; on final NLL the resolved losses are: none (FLCHAIN) and none (SUPPORT). **MIMIC-IV note.** For this learner, the MIMIC-IV comparison is inconclusive: follow-up is capped at day 365, and a probe at $k_{0.25}$ extends it by only a few percent of that horizon. After 400 probes, the fitted model changes little ($\mathrm{EPIG}_{\mathrm{surv}}$ 1.370→1.369; in a 40-seed pilot all ten rules lie within 0.0013 AULC of one another, $\mathrm{EPIG}_{\mathrm{surv}}$ ranked 2 behind C-BALD). Random probes with a larger budget, deeper windows or a smaller development set do not change this; only full outcomes would, which is a different problem. We therefore report MIMIC-IV for this learner as uninformative rather than as a win or a loss.

Table 11: Second learner (random survival forest), uniform costs, fresh seeds ($n = 40$). Mean paired difference (comparator − ours; positive favours ours) on AULC and on NLL at 400 probes, with its 95% bootstrap interval beneath and the verdict.

*Author-corrected total; one SUPPORT NLL row and its interval remain to be reconciled with this total.

| comparator | SUPPORT AULC | SUPPORT NLL@400 | FLCHAIN AULC | FLCHAIN NLL@400 |
|---|---|---|---|---|
| C-BALD | $+0.0154$ T<br>$[-0.0070, +0.0505]$ | $+0.0308$ T<br>$[-0.0000, +0.0882]$ | $-0.0013$ T<br>$[-0.0032, +0.0002]$ | $-0.0016$ T<br>$[-0.0045, +0.0014]$ |
| Entropy | $+0.0378$ W<br>$[+0.0073, +0.0746]$ | $+0.0678$ W<br>$[+0.0188, +0.1299]$ | $+0.0095$ W<br>$[+0.0042, +0.0154]$ | $+0.0126$ W<br>$[+0.0049, +0.0211]$ |
| Variance | $-0.0117$ T<br>$[-0.0405, +0.0068]$ | $+0.0059$ W<br>$[+0.0004, +0.0122]$ | $+0.0074$ W<br>$[+0.0015, +0.0141]$ | $+0.0090$ W<br>$[+0.0005, +0.0185]$ |
| IDEAL | $+0.0198$ W<br>$[+0.0021, +0.0437]$ | $+0.0223$ T<br>$[-0.0018, +0.0551]$ | $+0.0048$ W<br>$[+0.0004, +0.0098]$ | $+0.0066$ W<br>$[+0.0000, +0.0139]$ |
| Dedja | $+0.0373$ W<br>$[+0.0074, +0.0730]$ | $+0.0659$ W<br>$[+0.0186, +0.1245]$ | $+0.0086$ W<br>$[+0.0032, +0.0147]$ | $+0.0107$ W<br>$[+0.0032, +0.0191]$ |
| BatchBALD | $+0.0085$ T<br>$[-0.0172, +0.0324]$ | $+0.0246$ W<br>$[+0.0009, +0.0562]$ | $+0.0127$ W<br>$[+0.0069, +0.0192]$ | $+0.0197$ W<br>$[+0.0104, +0.0299]$ |
| Random | $-0.0216$ T<br>$[-0.0543, +0.0028]$ | $-0.0017$ T<br>$[-0.0064, +0.0030]$ | $+0.0063$ W<br>$[+0.0029, +0.0110]$ | $+0.0098$ W<br>$[+0.0026, +0.0183]$ |
| total (7 baselines) | 3W/4T/0L | 5W/2T/0L* | 6W/1T/0L | 6W/1T/0L |

Table 12: Second learner across the budget (fresh seeds, $n = 40$): $\mathrm{EPIG}_{\mathrm{surv}}$'s W/T/L against the seven baselines on NLL at the budget and on AULC up to it.

| probes | SUPPORT NLL | SUPPORT AULC | FLCHAIN NLL | FLCHAIN AULC |
|---|---|---|---|---|
| 100 | 2W/5T/0L | 2W/5T/0L | 3W/4T/0L | 4W/3T/0L |
| 200 | 4W/3T/0L | 3W/4T/0L | 6W/1T/0L | 6W/1T/0L |
| 300 | 3W/4T/0L | 3W/4T/0L | 6W/1T/0L | 6W/1T/0L |
| 400 | 5W/2T/0L | 3W/4T/0L | 6W/1T/0L | 6W/1T/0L |

# I SENSITIVITY TO THE NUMBER OF TIME BINS

**Number of time bins.** On SUPPORT (uniform costs, 40 development seeds) we re-ran $\mathrm{EPIG}_{\mathrm{surv}}$, C-BALD and Random with half (18) and double (69) the default 35 bins (Table 13). The two information-based rules remain close: $\mathrm{EPIG}_{\mathrm{surv}}$ has lower mean AULC at 18 and 69 bins, while C-BALD has a slightly lower mean at 35 bins. Both have lower mean AULC than Random at all three resolutions. NLL is not comparable across bin counts, so we compare methods only within each. This is a development-seed pilot, not a confirmation.

Table 13: SUPPORT, uniform costs, 40 development seeds: AULC of external NLL (lower is better, mean $\pm$ standard error over seeds) with the default number of time bins and with half and double; best value per column in bold. Compare within a column only.

| | 18 bins | 35 bins | 69 bins |
|---|---|---|---|
| $\mathrm{EPIG}_{\mathrm{surv}}$ (ours) | **2.0875**$\pm$0.0052 | 2.6346$\pm$0.0188 | **3.1422**$\pm$0.0162 |
| C-BALD | 2.0876$\pm$0.0044 | **2.6343**$\pm$0.0191 | 3.1462$\pm$0.0175 |
| Random | 2.1000$\pm$0.0046 | 2.6556$\pm$0.0178 | 3.1766$\pm$0.0226 |

# J MEAN ABSOLUTE ERROR WITH PSEUDO-OBSERVATIONS

**What MAE-PO is.** The most direct way to judge a survival model's time-to-event predictions is the mean absolute error (MAE) between the predicted and the true event times, but the true event time of a censored test patient is unknown. MAE with pseudo-observations (MAE-PO) (Qi et al., 2023a) fills this gap with a pseudo-observation. For each test patient $i$ the prediction $\hat{t}_i$ is the median of the model's predicted survival curve. An uncensored patient contributes the error $|t_i - \hat{t}_i|$. A censored patient contributes $|e_i - \hat{t}_i|$, where

$$e_i \;=\; N\,\hat{\theta} - (N-1)\,\hat{\theta}^{-i}$$

is the jackknife pseudo-observation of the Kaplan–Meier mean survival time (Andersen et al., 2003; Andersen & Pohar Perme, 2010): $\hat{\theta}$ is the Kaplan–Meier estimate of mean survival time on $N$ patients including $i$, and $\hat{\theta}^{-i}$ is the same estimate with $i$ left out. The pseudo-observation summarises patient $i$'s contribution to the estimated mean survival time; the construction used by Qi et al. (2023a) yields $e_i \geq t_i$ for a censored patient. It is a surrogate for the unobserved event time, not a recovered label. Each censored error is weighted by the confidence $w_i = 1 - S_{\mathrm{KM}}(t_i)$, which is low for a patient censored early, about whom little is known, and high for one censored late; uncensored patients have $w_i = 1$. Then

$$\text{MAE-PO} \;=\; \frac{1}{\sum_i w_i} \sum_i w_i \,\big|\,(1-\delta_i)\,e_i + \delta_i\,t_i - \hat{t}_i\,\big|,$$

reported in days; lower is better.

**Why MAE-PO.** MAE measures the absolute error of a time prediction in days. On uncensored data, expected absolute error is minimised by the conditional median (Qi et al., 2023a); this does not make MAE-PO a proper score for the full survival distribution. NLL, IBS and the C-index measure the likelihood, the probabilities over time and the ranking instead, so MAE-PO complements our primary metric. Among the ways of extending MAE to censored data — dropping the censored patients, hinge and margin losses, and two inverse-probability-of-censoring weightings — MAE-PO was the most reliable in the semi-synthetic benchmarks of Qi et al. (2023a), where the true event times of the censored patients are known: it identified the best-performing models and its value was often close to the true MAE, across several censoring distributions and censoring rates, including on MIMIC-IV. It also uses every censored test patient rather than discarding them, which matters on MIMIC-IV and FLCHAIN, where most patients are censored.

**Results.** At the final budget $\mathrm{EPIG}_{\mathrm{surv}}$ has no resolved loss against the seven baselines in the reported fixed-depth comparisons (Table 14): 48W/15T/0L over 63 comparisons on the development seeds, and 8W/6T/0L on the fresh uniform-cost seeds. Across the budget (Table 15) the few resolved losses occur before the final budget, on SUPPORT. Tables 16–17 give the values in days.

Table 14: MAE-PO at the final budget (400 probes with uniform costs; the common spend horizon with random and realised follow-up costs). Paired W/T/L of $\mathrm{EPIG}_{\mathrm{surv}}$ against the seven baselines. The predicted time is the median of each patient's predicted survival curve.

| setting | MIMIC-IV | SUPPORT | FLCHAIN |
|---|---|---|---|
| uniform costs | 6W/1T/0L | 3W/4T/0L | 6W/1T/0L |
| random costs | 6W/1T/0L | 6W/1T/0L | 7W/0T/0L |
| realised follow-up costs | 6W/1T/0L | 1W/6T/0L | 7W/0T/0L |
| uniform costs, fresh seeds | 6W/1T/0L | 2W/5T/0L | – |

Table 15: MAE-PO across the budget, development seeds. Each cell: W/T/L of $\mathrm{EPIG}_{\mathrm{surv}}$ against the seven baselines. With random and realised follow-up costs, checkpoints are fractions of the common spend horizon.

| setting | budget | MIMIC-IV | SUPPORT | FLCHAIN |
|---|---|---|---|---|
| uniform costs | 100 | 6W/1T/0L | 0W/7T/0L | 5W/2T/0L |
| | 200 | 6W/1T/0L | 0W/7T/0L | 6W/1T/0L |
| | 300 | 6W/1T/0L | 2W/4T/1L | 6W/1T/0L |
| | 400 | 6W/1T/0L | 3W/4T/0L | 6W/1T/0L |
| random costs | 25% | 6W/1T/0L | 3W/4T/0L | 7W/0T/0L |
| | 50% | 6W/1T/0L | 4W/3T/0L | 7W/0T/0L |
| | 75% | 7W/0T/0L | 3W/4T/0L | 7W/0T/0L |
| | 100% | 6W/1T/0L | 6W/1T/0L | 7W/0T/0L |
| realised follow-up costs | 25% | 6W/1T/0L | 0W/7T/0L | 6W/1T/0L |
| | 50% | 6W/1T/0L | 0W/7T/0L | 6W/1T/0L |
| | 75% | 6W/1T/0L | 1W/5T/1L | 7W/0T/0L |
| | 100% | 6W/1T/0L | 1W/6T/0L | 7W/0T/0L |

Table 16: MAE-PO in days at 400 probes, uniform costs, mean ± standard error over 40 development seeds (lower is better; bold = best in column).

| method | MIMIC-IV | SUPPORT | FLCHAIN |
|---|---|---|---|
| $\mathrm{EPIG}_{\mathrm{surv}}$ | **65.85**±0.60 | 596.15±4.71 | **705.93**±5.46 |
| C-BALD | 65.85±0.67 | 598.84±4.51 | 711.01±5.62 |
| Entropy | 67.39±0.56 | 613.43±5.29 | 725.43±6.17 |
| Variance | 67.25±0.54 | 600.39±3.99 | 722.48±6.00 |
| IDEAL | 67.18±0.56 | 603.07±4.63 | 716.36±5.81 |
| Dedja | 67.21±0.52 | 594.87±4.55 | 724.48±6.51 |
| BatchBALD | 68.06±0.59 | 601.43±5.81 | 720.78±6.39 |
| Random | 67.16±0.60 | **591.87**±5.70 | 712.24±6.32 |

Table 17: MAE-PO in days at the common spend horizon, random costs, mean ± standard error over 40 development seeds (lower is better; bold = best in column).

| method | MIMIC-IV | SUPPORT | FLCHAIN |
|---|---|---|---|
| $\mathrm{EPIG}_{\mathrm{surv}}$/cost | **65.87**±0.60 | **590.44**±4.57 | **702.24**±5.50 |
| C-BALD | 65.96±0.71 | 598.93±4.00 | 711.45±5.62 |
| Entropy | 67.51±0.56 | 613.17±5.21 | 725.68±6.16 |
| Variance | 67.21±0.54 | 600.45±4.18 | 723.13±5.99 |
| IDEAL | 67.25±0.50 | 603.05±4.72 | 716.70±5.85 |
| Dedja | 67.36±0.50 | 595.04±4.25 | 724.61±6.52 |
| BatchBALD | 68.05±0.60 | 601.26±6.02 | 721.26±6.37 |
| Random | 67.28±0.56 | 591.73±5.60 | 712.75±6.33 |

## K ADDITIONAL MATERIAL FOR THE FIXED-DEPTH SETTINGS

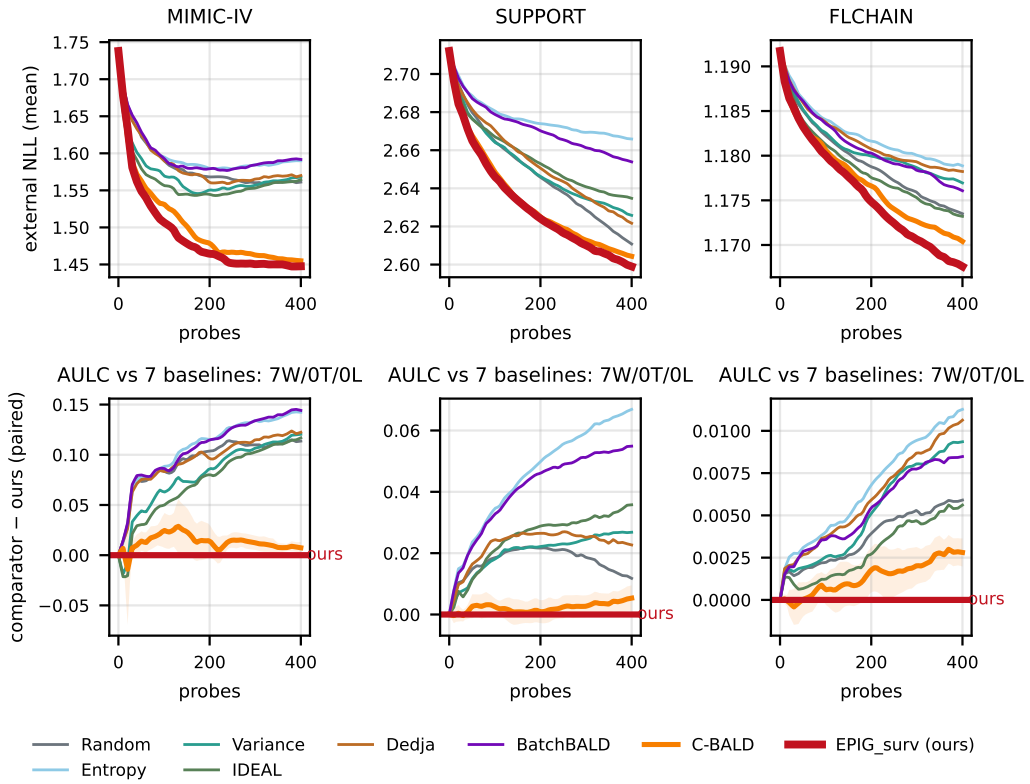



Figure 4: Uniform costs, all three datasets. Top: external censored NLL against probes bought, mean over seeds; ours is the thick red line. Bottom: each comparator's curve minus ours, paired by seed, with a 95% band for the nearest rival, C-BALD; above zero means ours is better. Panel titles give the AULC record against the seven baselines. $n = 40$ seeds.

Table 18: Uniform costs (development seeds, $n = 40$). Mean paired AULC difference (comparator − ours; positive favours ours) with its 95% bootstrap interval, and verdict; † survives the joint Holm correction.

| comparator | MIMIC-IV | SUPPORT | FLCHAIN |
|---|---|---|---|
| C-BALD | +0.0131 [+0.0057, +0.0202] W† | +0.0022 [+0.0002, +0.0047] W† | +0.0013 [+0.0007, +0.0020] W† |
| Entropy | +0.1076 [+0.0797, +0.1375] W† | +0.0456 [+0.0353, +0.0551] W† | +0.0067 [+0.0044, +0.0091] W† |
| Variance | +0.0800 [+0.0587, +0.1037] W† | +0.0203 [+0.0141, +0.0264] W† | +0.0052 [+0.0031, +0.0074] W† |
| IDEAL | +0.0727 [+0.0582, +0.0861] W† | +0.0248 [+0.0167, +0.0336] W† | +0.0029 [+0.0009, +0.0049] W† |
| Dedja | +0.0950 [+0.0729, +0.1167] W† | +0.0227 [+0.0139, +0.0312] W† | +0.0060 [+0.0038, +0.0084] W† |
| BatchBALD | +0.1074 [+0.0792, +0.1366] W† | +0.0407 [+0.0285, +0.0534] W† | +0.0054 [+0.0033, +0.0076] W† |
| Random | +0.0944 [+0.0699, +0.1195] W† | +0.0174 [+0.0092, +0.0247] W† | +0.0037 [+0.0014, +0.0060] W† |
| total (7 baselines) | 7W/0T/0L | 7W/0T/0L | 7W/0T/0L |

**Undefined response rows.** The response probabilities of Section 4.2 divide each draw's branch masses by their sum. When a draw's predictive pmf carries no mass after the candidate's current follow-up $t_i$ (numerical underflow of a member that is certain of an early event) the row is all zero after the division and is not a distribution; in the estimator such a draw contributes nothing to the response marginal and receives zero posterior weight, the resulting score no longer has the usual mutual-information interpretation. An audit of the saved round-0 states of the 40 pilot seeds, 7 candidate rows on MIMIC-IV and 0 elsewhere were affected, and 1 of these candidates was in a

Table 19: Random costs (development seeds, $n = 40$). Mean paired AULC difference on the common horizon (comparator − ours) with its 95% bootstrap interval, and verdict.

| comparator | MIMIC-IV | SUPPORT | FLCHAIN |
|---|---|---|---|
| C-BALD | $+0.0164$ $[+0.0085, +0.0251]$ W | $+0.0151$ $[+0.0091, +0.0218]$ W | $+0.0036$ $[+0.0026, +0.0047]$ W |
| Entropy | $+0.1106$ $[+0.0830, +0.1394]$ W | $+0.0584$ $[+0.0454, +0.0716]$ W | $+0.0088$ $[+0.0061, +0.0115]$ W |
| Variance | $+0.0835$ $[+0.0597, +0.1111]$ W | $+0.0336$ $[+0.0248, +0.0435]$ W | $+0.0074$ $[+0.0049, +0.0100]$ W |
| IDEAL | $+0.0757$ $[+0.0606, +0.0932]$ W | $+0.0380$ $[+0.0262, +0.0508]$ W | $+0.0053$ $[+0.0030, +0.0074]$ W |
| Dedja | $+0.0991$ $[+0.0760, +0.1214]$ W | $+0.0354$ $[+0.0246, +0.0464]$ W | $+0.0082$ $[+0.0056, +0.0108]$ W |
| BatchBALD | $+0.1108$ $[+0.0818, +0.1396]$ W | $+0.0534$ $[+0.0389, +0.0682]$ W | $+0.0076$ $[+0.0051, +0.0100]$ W |
| Random | $+0.0974$ $[+0.0731, +0.1211]$ W | $+0.0304$ $[+0.0207, +0.0415]$ W | $+0.0060$ $[+0.0031, +0.0087]$ W |
| total (7 baselines) | 7W/0T/0L | 7W/0T/0L | 7W/0T/0L |

Table 20: Uniform costs, fresh-seed confirmation ($n = 40$, protocol identical to the pilot). Mean paired difference (comparator − ours) on AULC and on NLL at 400 probes, with its 95% bootstrap interval beneath and the verdict; totals over the seven baselines.

| comparator | MIMIC-IV AULC | MIMIC-IV NLL@400 | SUPPORT AULC | SUPPORT NLL@400 |
|---|---|---|---|---|
| C-BALD | $+0.0045$ T | $+0.0026$ T | $+0.0027$ W | $+0.0002$ T |
| | $[-0.0013, +0.0121]$ | $[-0.0024, +0.0076]$ | $[+0.0004, +0.0049]$ | $[-0.0024, +0.0030]$ |
| Entropy | $+0.1074$ W | $+0.1381$ W | $+0.0369$ W | $+0.0553$ W |
| | $[+0.0824, +0.1329]$ | $[+0.1036, +0.1742]$ | $[+0.0215, +0.0494]$ | $[+0.0233, +0.0792]$ |
| Variance | $+0.0894$ W | $+0.1317$ W | $+0.0149$ W | $+0.0177$ T |
| | $[+0.0714, +0.1064]$ | $[+0.1074, +0.1572]$ | $[+0.0013, +0.0257]$ | $[-0.0026, +0.0345]$ |
| IDEAL | $+0.0864$ W | $+0.1224$ W | $+0.0255$ W | $+0.0428$ W |
| | $[+0.0685, +0.1034]$ | $[+0.0948, +0.1500]$ | $[+0.0116, +0.0371]$ | $[+0.0194, +0.0641]$ |
| Dedja | $+0.0948$ W | $+0.1196$ W | $+0.0166$ W | $+0.0183$ T |
| | $[+0.0749, +0.1129]$ | $[+0.0925, +0.1481]$ | $[+0.0007, +0.0308]$ | $[-0.0104, +0.0436]$ |
| BatchBALD | $+0.1056$ W | $+0.1336$ W | $+0.0337$ W | $+0.0524$ W |
| | $[+0.0804, +0.1317]$ | $[+0.0989, +0.1690]$ | $[+0.0156, +0.0479]$ | $[+0.0206, +0.0780]$ |
| Random | $+0.0948$ W | $+0.1143$ W | $+0.0160$ T | $+0.0179$ T |
| | $[+0.0736, +0.1151]$ | $[+0.0847, +0.1451]$ | $[-0.0031, +0.0311]$ | $[-0.0174, +0.0445]$ |
| total (7 baselines) | 6W/1T/0L | 6W/1T/0L | 6W/1T/0L | 3W/4T/0L |

recorded first batch (seed 23010, sixteen of sixty-four draws, i.e. one ensemble member). Intermediate states were not kept, so later rounds cannot be re-read; the frozen harness was not changed.

**Two dataset-specific mechanisms (full account of Section 6.1).** On MIMIC-IV the non-random methods that lose most (Entropy, Variance, IDEAL, Dedja, BatchBALD) query the fewest distinct patients and re-probe the most (160–186 distinct patients out of 400 probes, versus 235–239 for C-BALD and ours): their scores keep re-selecting patients whose window keeps returning "no endpoint yet". This is one mechanism, not a universal one (see the selection-noise paragraph below). On FLCHAIN (Figure 7b) the label-informed oracle, which maximises revealed events, ends *worse* than the refit-only control (final NLL 1.202 vs 1.181; every outcome-blind method 1.168–1.179). This is informative censoring introduced by the policy, not covariate shift (every pool patient is in the training set; the oracle only de-censors): extending follow-up precisely for patients known to die inside the window makes whether a patient stays censored depend on $T$, violating $T \perp C \mid X$ and inflating the fitted in-window hazard. The oracle's event NLL improves while its censored NLL deteriorates (Appendix K): the model becomes pessimistic on a 72%-censored test population. Event count alone does not explain the observed ordering: among the 8 outcome-blind methods revealed events and final NLL correlate at $-0.06$ (Figure 7b). Under the stated censoring assumptions, restricting acquisition to observed history avoids this direct use of hidden outcomes. The label-informed oracle therefore provides a reference rather than a performance upper bound.

**FLCHAIN oracle decomposition.** The oracle's event NLL improves from 3.53 to 3.42 while its censored NLL deteriorates from 0.30 to 0.35 (every outcome-blind method ends at 0.24–0.27 on the censored term); across all 10 methods, including the two near-identical oracle points, the correlation between events revealed and final NLL is $+0.78$. The refit-only control accounts for 49/28/45% (MIMIC-IV/SUPPORT/FLCHAIN) of $\mathrm{EPIG}_{\mathrm{surv}}$'s initial-to-final improvement with uniform costs (Figure 6); every arm has the same refitting schedule, so the schedule is matched across arms, though

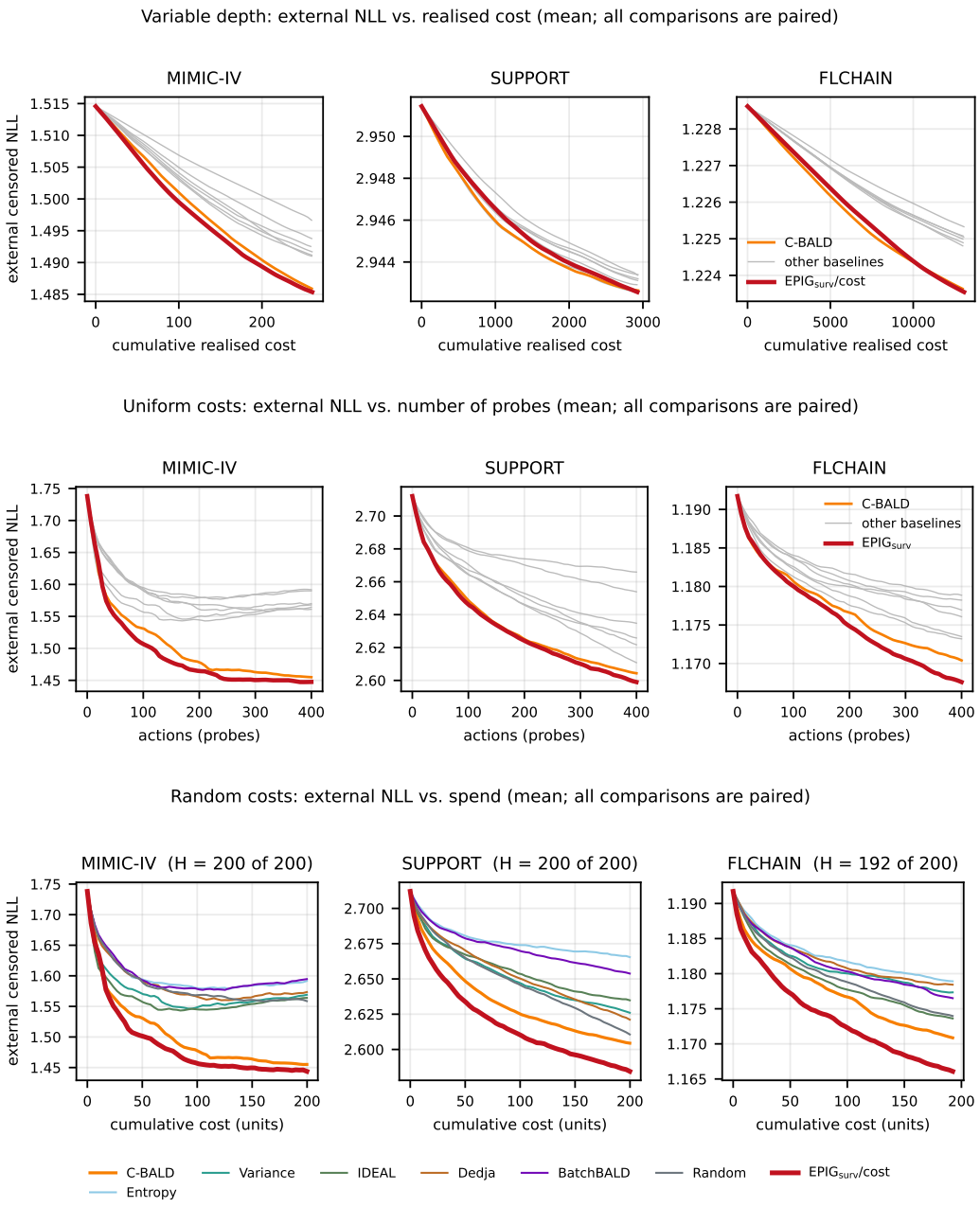



Figure 5: Learning curves (mean over seeds) with variable depth (top; realised cost axis), uniform costs (middle; probes) and random costs (bottom; cost units, common horizon $H$). Absolute performance varies across data splits, so all method comparisons are paired by seed. Only ours and the two nearest rivals are coloured. $n = 40$ seeds.

it can interact differently with different acquired data; absolute gains should therefore be interpreted alongside the refit-only control.

**Selection noise.** In a diagnostic on 40 development seeds per dataset with the round-0 model frozen, an eight-member ensemble was fitted so that the production posterior (members 1–4, i.e. the first 64 of the 128 draws, exactly what a four-member run uses) can be compared with a disjoint one (members 5–8). Re-scoring with the disjoint members substantially changes the selected batch (Table 21: top-10 Jaccard against the production batch 0.14–0.32 for $\mathrm{EPIG}_{\mathrm{surv}}$, 0.16–0.43 for C-BALD; score

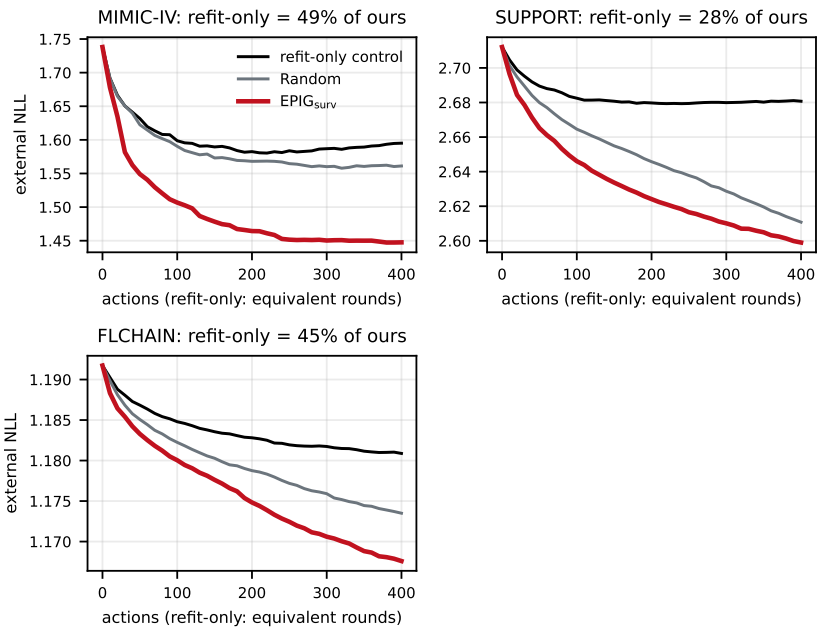
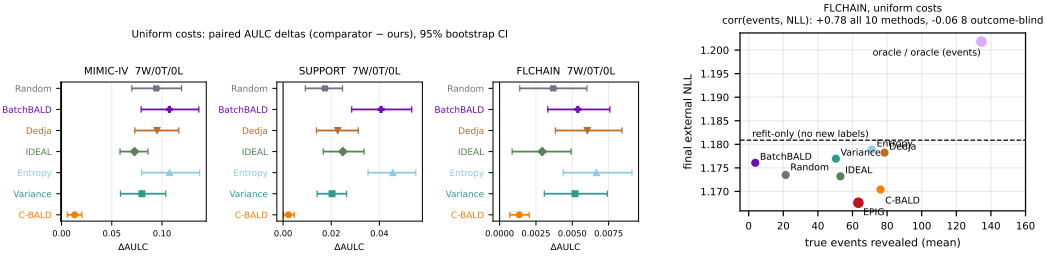



Figure 6: Optimisation drift control with uniform costs on all three datasets: black = refit-only control (no new labels, same re-fitting schedule), grey = Random, red = $\mathrm{EPIG}_{\mathrm{surv}}$. Whether an unconverged initial fit also interacts with the acquired data cannot be tested here, and no refit-only arm was run in the variable-depth setting. $n = 40$ seeds.




(a) Uniform costs: paired AULC deltas.

(b) FLCHAIN, uniform costs.

Figure 7: (a) Paired AULC deltas against $\mathrm{EPIG}_{\mathrm{surv}}$ with 95% intervals; panel titles count the seven baselines. (b) The label-informed oracle ends above the no-label control; among the 8 outcome-blind methods the observed correlation between revealed events and final NLL is small ($-0.06$). $n = 40$ seeds.

Spearman 0.48–0.77), whereas cutting the target sample from 300 to 75 barely changes it (Jaccard 0.89–1.00, Spearman 1.00–1.00). Among the perturbations tested, ensemble composition is the larger source of selection variability. An earlier version of this diagnostic sliced posterior draws rather than members and understated the effect. $\mathrm{EPIG}_{\mathrm{surv}}$ and C-BALD select overlapping batches (Jaccard 0.43–0.83; FLCHAIN 0.43). On SUPPORT the re-probing mechanism does not hold: Entropy and Dedja query 359 and 380 distinct patients, and nevertheless underperform $\mathrm{EPIG}_{\mathrm{surv}}$. On FLCHAIN the uniform-cost probe depth is 559–1124 days across seeds against a median observed time of 4,302 days and a Kaplan–Meier median that is not reached within follow-up; $\mathrm{EPIG}_{\mathrm{surv}}$ reveals 63 events, the most natural censorings of any method (114) and, Random aside, the most distinct patients (273; Random 374).

**Random-cost invariants.** The recorded cost of every action matches the patient-specific draw from $U(0.2, 0.8)$ generated with random seed $\mathrm{seed} + 4001$; a patient re-probed in a later round is charged

Table 21: Selection-noise diagnostic with uniform costs (40 development seeds per dataset, round-0 posterior frozen; mean ± standard error over seeds). Top-10 Jaccard and score Spearman of each perturbed scorer against the production one: target sample cut to 75, and a disjoint four-member ensemble.

| dataset | targets 300 → 75 | | disjoint 4 members, $\mathrm{EPIG}_{\mathrm{surv}}$ | | disjoint 4 members, C-BALD | | $\mathrm{EPIG}_{\mathrm{surv}}$ vs C-BALD |
|---|---|---|---|---|---|---|---|
| | Jaccard | Spearman | Jaccard | Spearman | Jaccard | Spearman | Jaccard |
| MIMIC-IV | 0.94±0.06 | 1.00±0.00 | 0.32±0.11 | 0.70±0.02 | 0.43±0.06 | 0.76±0.01 | 0.78±0.11 |
| SUPPORT | 1.00±0.00 | 1.00±0.00 | 0.14±0.07 | 0.48±0.03 | 0.16±0.06 | 0.56±0.04 | 0.83±0.10 |
| FLCHAIN | 0.89±0.11 | 1.00±0.00 | 0.17±0.09 | 0.77±0.02 | 0.31±0.06 | 0.87±0.02 | 0.43±0.06 |

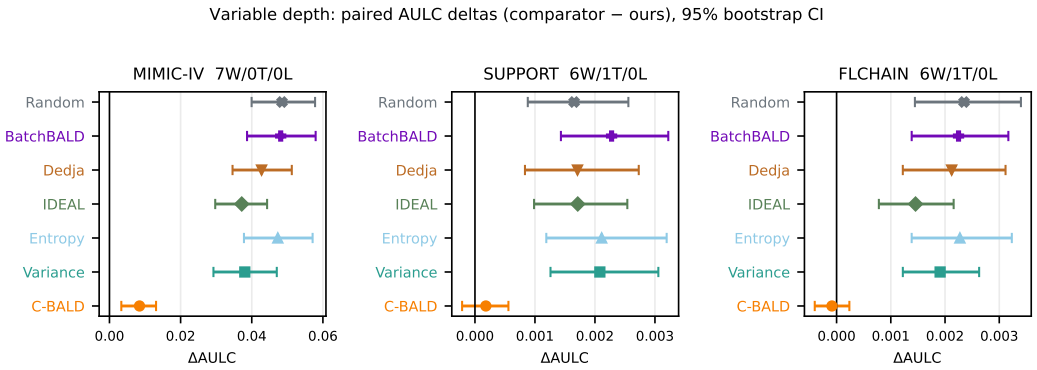



Figure 8: Variable depth: paired AULC deltas against $\mathrm{EPIG}_{\mathrm{surv}}$/cost with 95% bootstrap intervals; panel titles count the seven baselines; one colour and marker per method, as in Figure 1. $n = 40$ seeds.

the identical cost; costs are assigned independently of outcomes; and no cost lies outside $[0.2, 0.8]$. On FLCHAIN the common horizon is $H = 192.35$ of 200: it was set by an additional arm run in the same experiment and not reported here, whose greedy selector returned partial batches in three seeds and reached the implementation's 60-round limit.

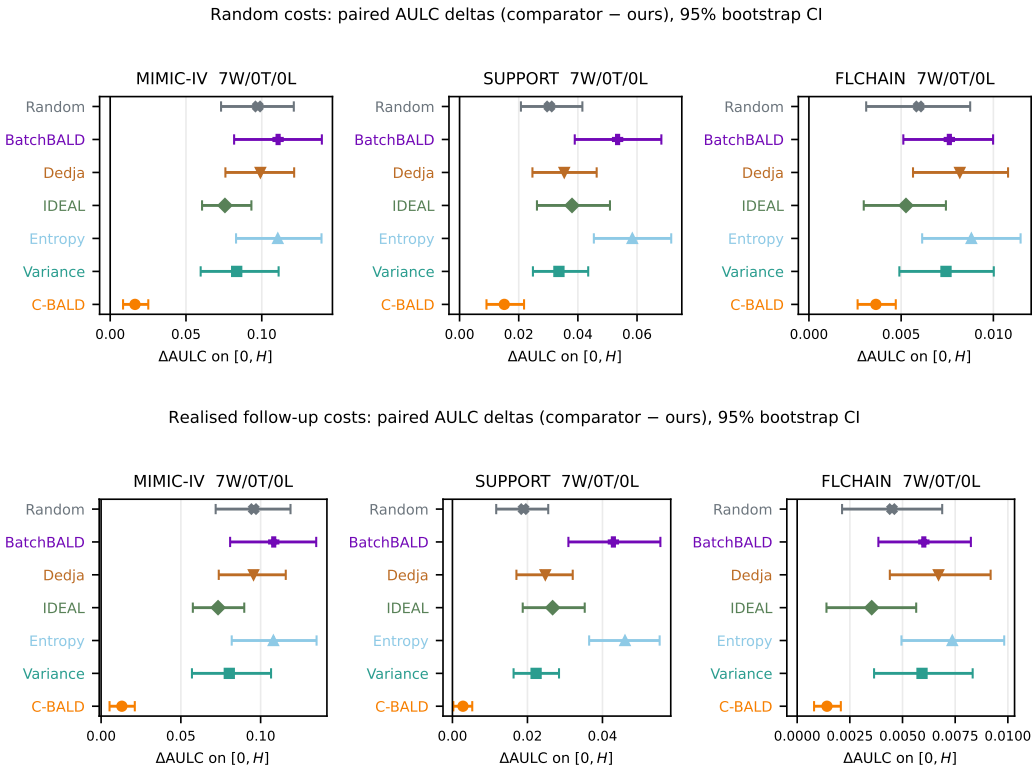



Figure 9: Random and realised follow-up costs: paired AULC deltas on the common horizon; panel titles count the seven baselines. Top: random costs. Bottom: realised follow-up costs, where each seed's cost axis is normalised by its own budget $575\,k_{0.25}$ because $k_{0.25}$ is a data-derived quantile that differs across seeds. $n = 40$ seeds.