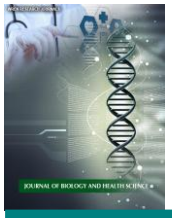

Journal of Biology and Health Science

ISSN: 3108-0189

Research Article

WRJBHS-26-152

Selection-Bias-Aware Multimodal Survival Learning for Molecular Pathological Epidemiology: A Stabilized IPW–Cox Method with a Differentiable Etiologic-Heterogeneity Head

Young Sahng Suh[1*] **and Shuji Ogino**[2]

[1]aSSIST University, Seoul, Republic of Korea
[2]Program in Molecular Pathological Epidemiology (MPE), Department of Pathology, Brigham and Women's Hospital and Harvard Medical School, Boston, USA

[*]**Correspondence:** Young Sahng Suh, aSSIST University, Seoul, Republic of Korea, E-mail: yssuh@themachinaai.com; DOI: https://doi.org/10.56147/jbhs.3.3.152

Citation: Suh YS, Ogino S (2026) Selection-Bias-Aware Multimodal Survival Learning for Molecular Pathological Epidemiology: A Stabilized IPW–Cox Method with a Differentiable Etiologic-Heterogeneity Head. J Biol & Heal Sci 3: 152.

Abstract

Background: Deep survival models for tumor-tissue and molecular studies are trained only on patients whose specimens are collected, archived, retrieved, assayed and pass quality control. In Molecular Pathological Epidemiology (MPE) this analyzable subset is a covariate-dependent draw from the incident-case population, because tissue retrieval, assay success and molecular-subtype ascertainment all depend on prognostic variables. A model that ignores this targets the analyzable-sample relationship rather than the population one.

Methods: We developed SBASURV, an architecture-agnostic survival head that (1) replaces the ordinary Cox partial likelihood with a stabilized Inverse-Probability-Weighted (IPW) Cox objective in which availability weights enter both the event term and the risk-set denominator and (2) implements the Lunn-McNeil duplication method as a differentiable subtype-specific module emitting exposure-by-subtype log-hazard ratios and a Wald statistic for etiologic heterogeneity. We validated the mechanism on five public survival datasets with induced covariate-dependent selection, then deployed the identical head on TCGA-COAD/READ colorectal cancer across three modalities clinical, tumor omics and whole-slide histopathology (features from 441 diagnostic slides *via* a public foundation model) using real tissue analyzability weights.

Findings: The differentiable stabilized-IPW head improved held-out concordance over a naive neural head on all five public datasets, while weight diagnostics made the bias-variance regime explicit. On real colorectal cancer, clinical+omics fusion reached held-out Harrell C 0.68 (Uno 0.66) in 530 cases; adding pathology improved discrimination over clinical alone (0.59 to 0.62) in the 367-case three-modality sub cohort. Availability weighting preserved discrimination while shifting the estimand with well-conditioned weights (mean ≈ 1, effective sample size 96–99%). A tumor-mutation-burden/MSI proxy subtype yielded significant stage-by-subtype heterogeneity (Q=10·2, p=0·001), but a genuine methylation-defined CIMP subtype did not (Q=0·01, p=0·92).

Interpretation: SBASURV provides a principled, differentiable mechanism for population-targeted multimodal survival learning. The contrast between proxy and assay-grounded subtypes shows that built-in heterogeneity tests are valuable as much for the nulls they return as for the signals and must be read with weight and sample-size diagnostics in view. The method is a research instrument, not a clinical tool.

Keywords: Selection bias; Inverse probability weighting; Survival analysis; Molecular pathological epidemiology; Multimodal machine learning; Etiologic heterogeneity; NHANES; Colorectal cancer; Computational pathology; Digital health

Received date: May 27, 2026; **Accepted date:** June 04, 2026; **Published date:** June 17, 2026

Journal of Biology and Health Science

ISSN: 3108-0189

Introduction

Survival learning is increasingly used for clinical risk modeling, treatment stratification and prognostic discovery. Building on the Cox proportional-hazards model, neural and ensemble survival models improve time-to-event prediction when training and deployment distributions are aligned [1]. A weaker assumption is usually left implicit: That the analyzable training sample represents the population to which the learned relationship will be applied [2-8]. This assumption is fragile in tumor-tissue studies. A Colorectal-Cancer (CRC) model that uses whole-slide images, tumor multiomics or tissue-derived molecular sub-types can only be trained on patients whose tumor material was collected, archived, retrieved, assayed and passed quality control. Each step is a selection process and availability can depend on stage, site, age, comorbidity, institution, surgical pathway, specimen age and follow-up intensity variables that are also prognostic. The prospective-cohort incident-tumor-biobank framework underpinning modern MPE makes this selection explicit and motivates correcting for it directly [9-12]. Epidemiology already has the language for the problem. Inverse-probability weighting traces to Horvitz-Thompson estimation and the propensity score, underlies estimating-equation correction for missing and selection and in its stabilized form corrects estimating equations when selection probabilities can be modeled from observed covariates; robust and weighted Cox variance estimators have long precedent [13-21]. In MPE, IPW has handled missing tumor features and biospecimen availability [22]. The Lunn-McNeil duplication method supports subtype-specific Cox modeling and etiologic-heterogeneity tests ask whether exposure-outcome associations differ across molecular subtypes [23-24]. On the imaging side, histopathology foundation models 25–30 and weakly supervised slide-level learning now predict molecular status and prognosis from tissue and multimodal histology–genomic fusion has improved cancer prognosis [31-36]. These tools are standard individually but have not been built into the trainable objective of a selection-bias-aware multimodal survival learner [37-38]. That is the gap we close.

Methods

Availability model and stabilized weights

For individual i we observe $O_i = (T_i, \Delta_i, X_i, S_i, A_i)$ with follow-up T_i, event indicator Δ_i, covariates or learned representation X_i, molecular subtype S_i and analyzability indicator $A_i \in \{0, 1\}$. The availability propensity is $\hat{\pi}(x) = \hat{P}(A = 1 \mid X = x)$, estimated from variables observed for both analyzable and non-analyzable individuals. With $\hat{P}_A = n^{-1} \sum_i A_i$, the stabilized weight is $sw_i = \hat{P}_A / \hat{\pi}(x_i)$, clipped at the 1st/99th empirical percentiles. We report min, percentiles, mean, SD, max and effective sample size ESS = $(\Sigma_i \bar{w}_i)^2 / \Sigma_i \bar{w}_i^2$ with every result.

Stabilized-IPW Cox partial likelihood

For mini-batch B with in-batch risk set RB(Ti), the weighted negative Cox partial likelihood is

$$q_{C_0 x}(\theta) = -\sum_{i \in B} \Delta_i \bar{w}_i \left[f_\theta(h_i) - log \sum_{j \in RB(Ti)} \bar{w}_j \, exp\left(F_\theta(h_j)\right) \right] \ldots (1)$$

The risk-set denominator is weighted because the risk set is itself sampled through the availability process; setting all $\bar{w}_i$=1 recovers the standard neural Cox objective. The denominator uses a log-sum-exp form for numerical stability.

Differentiable duplication-method heterogeneity head

With exposure Ei and subtype-interacted log-risk $f_{\theta,\beta}(h_i, k) = g_\theta(hi) + \beta_k E_i$, the Lunn–McNeil construction is a stratified weighted Cox objective with subtype-specific event indicators and exposure columns. Coefficients β are tested under H_0: $\beta_1 = \cdots = \beta_K$ with a successive-difference contrast C and robust co-variance $\hat{\Sigma}_\beta$: $Q = \left(c\hat{\beta}\right)^T \left(c\hat{\Sigma}_\beta c^T\right)^{-1} \left(c\hat{\beta}\right) \sim \chi^2_{k-1}$. By default, Q is an inferential output, not a training penalty.

Multimodal encoders and fusion

The head is modality-agnostic. In the CRC deployment we use (1) clinical/demographic features; (2) tumor omics (RNA-seq PCs, tumor mutation burden, hypermutation); and (3) whole-slide histopathology foreground tiles at ~20 × embedded with a public histopathology foundation model and mean-pooled to a slide vector reduced to an unsupervised principal-component block. This label-blind aggregation is a conservative lower bound on the pathology contribution.

Statistical theory

Under missing-at-random availability, positivity and Cox regularity, the stabilized-IPW Cox score computed in the analyzable sample shares the population zero of the full-cohort Cox score, so $\hat{\theta}$ is consistent for θ_0 and asymptotically normal with a robust sandwich variance. With clipped weights $\bar{w}_i \in$ [a, b], the mini-batch gradient second moment is bounded by b^2G^2/m, explaining why clipping is necessary for stable optimization.

Findings

Mechanism on public benchmarks

On five public datasets (colon, GBSG2, SUPPORT, METABRIC, gbsg; assembled *via* SurvSet and the DeepSurv benchmarks 2) with induced covariate-dependent selection, the differentiable stabilized-IPW head improved held-out Harrell C over a naive neural head on all five (eg

Journal of Biology and Health Science

ISSN: 3108-0189

gbsg 0·609 to 0·639; colon 0·599 to 0·628; SUPPORT 0·505 to 0·539) [39-41]. Classical post-hoc weighting gave a deliberately mixed coefficient-recovery picture: Stabilized IPW improved RMSE on METABRIC and gbsg but worsened it where overlap was weak, the expected bias-variance signature of IPW. Stabilized weights had mean ≈ 1·0 and clipped-weight ESS of roughly 0·65-0·80n; uncapped maxima reached ≈ 27 on SUPPORT, confirming why clipping is required **(Tables 1 and 2; Figures 1-3)**.

Table 1: Classical post-hoc weighted Cox under induced covariate-dependent selection (b=1·3, mean ± SD across 20 selection draws). C-index on a representative held-out test set; coefficient RMSE relative to the oracle Cox fit.

Dataset	n	Oracle C	Naive C	Stabilized-IPW C
Colon	929	0·649	0·637 ± 0·007	0·637 ± 0·012
GBSG2	686	0·630	0·624 ± 0·007	0·620 ± 0·010
Support	8873	0·572	0·573 ± 0·002	0·570 ± 0·003
Metabric	1904	0·645	0·638 ± 0·003	0·636 ± 0·006
gbsg	2232	0·650	0·639 ± 0·005	0·644 ± 0·006

Table 2: Differentiable neural stabilized-IPW Cox. Held-out Harrell C, mean ± SD over 10 induced-selection draws; ESS frac is the mean clipped-weight ESS as a fraction of analyzable n.

Dataset	n	Oracle C	Neural naive C	Neural stab-IPW C	ESS frac
Colon	929	0·602	0·599 ± 0·005	0·628 ± 0·013	0·75
GBSG2	686	0·589	0·603 ± 0·026	0·624 ± 0·031	0·78
Support	8873	0·480	0·505 ± 0·018	0·539 ± 0·014	0·65
Metabric	1904	0·597	0·579 ± 0·029	0·600 ± 0·009	0·65
gbsg	2232	0·612	0·609 ± 0·006	0·639 ± 0·010	0·80

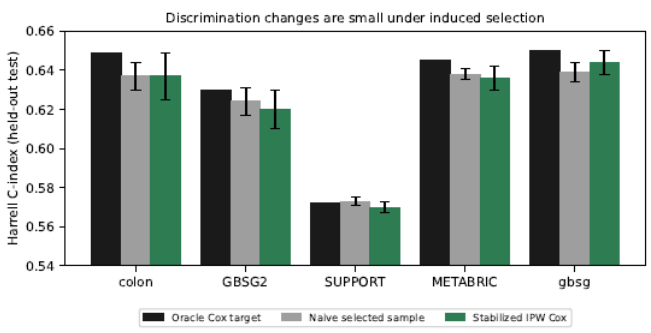

Figure 1: C-index from the post-hoc correction; the correction does not produce a uniform discrimination change.

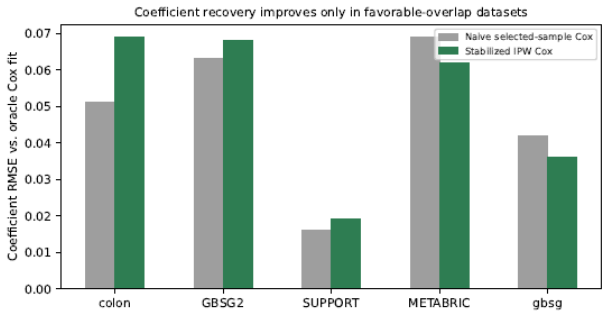

Figure 2: Coefficient recovery relative to the oracle. Stabilized IPW improves RMSE on METABRIC and gbsg but not on colon, GBSG2 or SUPPORT.

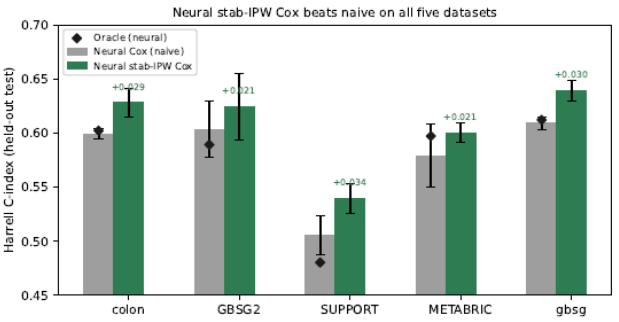

Figure 3: The differentiable neural stabilized-IPW Cox head improves held-out C-index over the naive neural head on all five datasets (diamonds=oracle).

Real TCGA-CRC deployment

We assembled 633 TCGA-COAD/READ cases. The analyzable indicator was 1 for usable RNA-seq and survival follow-up (n=530, 122 events) [42]. The propensity was genuinely covariate-dependent (AUC 0·78) yet satisfied positivity; stabilized clipped weights were well-conditioned (mean 0·999, max 1·92, ESS 510/530, 96%). Clinical+omics fusion reached Harrell 0·68 (Uno C 0·66); availability weighting left fusion discrimination essentially unchanged (0·679-0·680) while shifting coefficients ($||\Delta\beta||_2$=0·014) **(Table 3 and Figure 4)** [43].

Table 3: Real TCGA-COAD/READ, omics+clinical. Held-out C-index (mean±SD over repeated 5-fold CV; n=530, 122 events) for naive *vs.* availability-weighted SBASURV heads, by modality ablation.

Feature block	d	Harrell (naive)	Harrell (weighted)	Uno (weighted)
Omics-only (RNA-seq PCs)	16	0·604 ± 0·061	0·566 ± 0·080	0·556 ± 0·077
Clinical-only	6	0·607 ± 0·055	0·576 ± 0·057	0·567 ± 0·064
Clinical+omics	22	0·679 ± 0·059	0·680 ± 0·052	0·659 ± 0·052

Journal of Biology and Health Science

ISSN: 3108-0189

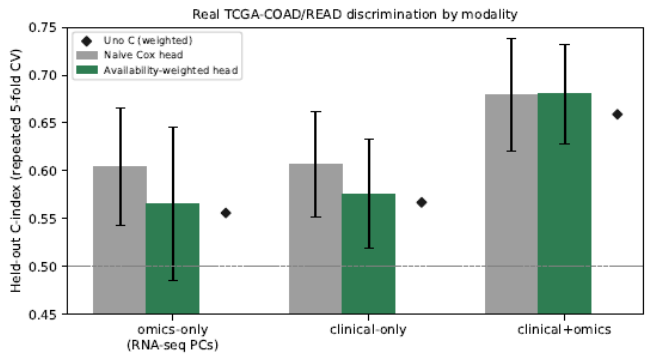

Figure 4: Real TCGA-COAD/READ discrimination by modality ablation. Clinical+omics fusion beats single modalities and availability weighting preserves discrimination while changing the estimand.

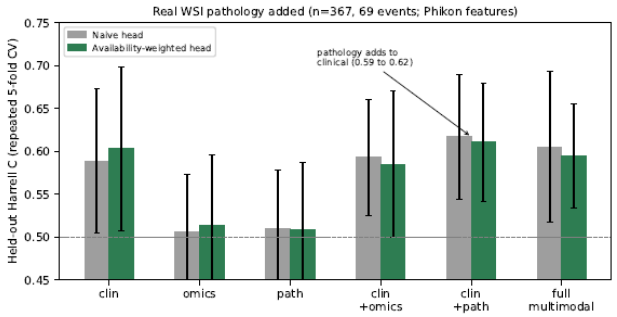

Figure 5: Adding real whole-slide pathology (Phikon features) to the TCGA-CRC head: on the three-modality cohort pathology improves discrimination over clinical alone, with availability weighting pre-serving the ranking.

Whole-slide pathology and heterogeneity

Pathology features were extracted from 441/442 diagnostic slides using a public histopathology foundation 120 model. On the three-modality cohort (n=367, 69 events) pathology added complementary signal: Clinical +pathology (0·617) was best and improved on clinical alone (0·589) [26,30]. The heterogeneity head exposed a key MPE caution: A TMB/MSI proxy subtype gave significant stage-by-subtype heterogeneity (availability-weighted Q=10·2, p=0·0014; n=439), but a genuine five-gene-methylation CIMP subtype did not (Q=0·01, p=0·92; n=296) **(Table 4; Figure 5 and 6)** [34,44].

Table 4: Real TCGA-COAD/READ, three modalities with whole-slide pathology (Phikon features). Held-out C-index (mean±SD, repeated 5-fold CV) on the three-modality analyzable cohort (n=367, 69 events).

Feature block	d	Harrell (naive)	Harrell (weighted)	Uno (weighted)
Clinical-only	6	0·589 ± 0·084	0·603 ± 0·095	0·614 ± 0·117
Omics-only (RNA-seq PCs)	16	0·506 ± 0·067	0·514 ± 0·082	0·520 ± 0·086
Pathology-only (Phikon PCs)	16	0·510 ± 0·068	0·509 ± 0·078	0·542 ± 0·078
Clinical+omics	22	0·593 ± 0·068	0·585 ± 0·085	0·625 ± 0·068
Clinical+pathology	22	0·617 ± 0·073	0·611 ± 0·069	0·637 ± 0·078
Omics+pathology	32	0·499 ± 0·078	0·509 ± 0·092	0·546 ± 0·089
Clinical+omics+pathology	38	0·605 ± 0·088	0·595 ± 0·061	0·631 ± 0·074

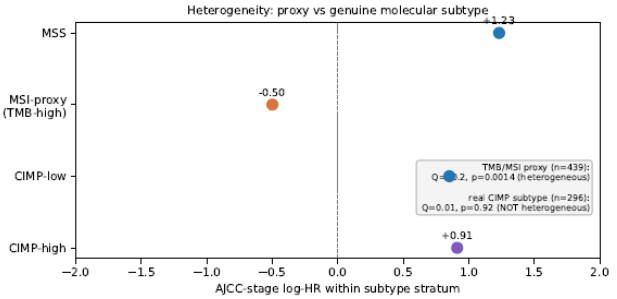

Figure 6: Heterogeneity of the AJCC-stage effect by molecular subtype: A TMB/MSI proxy shows significant heterogeneity (Q=10·2, p=0·001), but a genuine CIMP epigenetic subtype does not (Q=0·01, p=0·92).

Discussion

SBASURV treats sample availability as a first-class statistical object in survival learning. Two deployment findings resist hype: Real availability weighting changed the estimand while leaving discrimination essentially intact the correct behavior for a population-targeting method, since concordance is not the metric IPW is designed to move; and the heterogeneity machinery returned a significant result on a proxy subtype yet a null on a genuine one, demonstrating that a built-in test is valuable as much for its nulls as its signals. The pathology result is deliberately modest: A mean-pooled, label-blind embedding contributes complementary but not stand-alone prognostic signal at this event count, consistent with the computational-pathology literature where strong slide-level prognosis generally requires trained multiple-instance aggregation [31,32].

All real results are within TCGA; generalization to prospective, multi-institution and non-Western cohorts is untested and consensus molecular subtypes could not be called from the assembled feature set [45]. Consistency rests on missing-at-random availability, positivity and a correctly specified propensity model; residual selection on

unmeasured prognostic variables would not be removed by weighting on observed covariates. The model is a research instrument for population-targeted survival learning and etiologic-heterogeneity testing, not a clinical tool; calibration, integrated Brier score, decision-curve analysis, subgroup fairness and external replication reported following TRIPOD+AI are the prerequisites for any clinical reading [46-48].

Contributors

YSS conceived the study, developed the method and theory, implemented all experiments, analyzed the data and wrote the manuscript. YSS had full access to all data and final responsibility for the decision to submit.

Declaration of Interests

The author declares no competing interests.

Data Sharing

All public datasets are openly available (DeepSurv benchmarks, SurvSet, R survival: Colon). TCGA-COAD/READ clinical, RNA-seq, somatic-mutation, methylation and diagnostic whole-slide images are available from the NIH Genomic Data Commons. The full pipeline, availability-model specifications, per-run weight diagnostics, seeds and figure scripts are released to regenerate every table and figure.

Use of AI Tools

AI-assisted software tools were used to help implement experiment, pathology and figure-generation code and to assist in correcting minor errors in text; all reported numbers are genuine reproducible outputs of the released code on real public and GDC data and the author is responsible for every claim.

Funding

None.

Role of the Funding Source

There was no funding source. The author had full access to all data and final responsibility for the decision to submit.

References

1. Cox DR (1972) Regression models and life-tables. Journal of the Royal Statistical Society: Series B 34: 187-202. [Crossref] [Google Scholar]
2. Jared LK, Uri S, Alexander C, Jonathan B, Tingting J, et al. (2018) DeepSurv: Personalized treatment recommender system using a Cox proportional hazards deep neural network. BMC Medical Research Methodology 18: 24. [Crossref] [Google Scholar] [Indexed]
3. Ching T, Zhu X, Lana XG (2018) Coxnnet: An artificial neural network method for prognosis prediction of high-throughput omics data. PLoS Computational Biology 14: e1006076. [Crossref] [Google Scholar] [Indexed]
4. Lee C, William RZ, Jinsung Y, Mihaela S (2018) DeepHit: A deep learning approach to survival analysis with competing risks. AAAI Conference on Artificial Intelligence 32: 2314-2321. [Crossref] [Google Scholar]
5. Kvamme H, Ørnulf B, Scheel I (2019) Time-to-event prediction with neural networks and Cox regression. Journal of Machine Learning Research 20: 1-30. [Google Scholar]
6. Wang Z, Sun J (2022) SurvTRACE: Transformers for survival analysis with competing events. 13th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics: 1-9. [Crossref] [Google Scholar]
7. Ishwaran H, Kogalur UB, Blackstone EH, Lauer MS (2008) Random survival forests. Annals of Applied Statistics 2: 841-860. [Crossref] [Google Scholar]
8. Gensheimer MF, Narasimhan B (2019) A scalable discrete-time survival model for neural networks. PeerJ 7: e6257. [Crossref] [Google Scholar] [Indexed]
9. Ogino S, Stampfer MJ (2010) Lifestyle factors and microsatellite instability in colorectal cancer: The evolving field of molecular pathological epidemiology. Journal of the National Cancer Institute 102: 365-367. [Crossref] [Google Scholar] [Indexed]
10. Ogino S, Chan AT, Fuchs CS, Giovannucci E (2011) Molecular pathological epidemiology of colorectal neoplasia: An emerging transdisciplinary and interdisciplinary field. Gut 60: 397-411. [Crossref] [Google Scholar] [Indexed]
11. Ogino S, Lochhead P, Chan AT, Nishihara R, Cho E, et al. (2013) Molecular pathological epidemiology of epigenetics: Emerging integrative science to analyze environment, host and disease. Modern Pathology 26: 465-484. [Crossref] [Google Scholar] [Indexed]
12. Ogino S, Ugai S, Hamada T, Ugai T (2025) The early-onset cancer epidemics: Evidence synthesis using the prospective cohort incident-tumor biobank method. European Journal of Epidemiology 40: 1405-1417. [Crossref] [Google Scholar] [Indexed]
13. Horvitz DG, Thompson DJ (1952) A generalization of sampling without replacement from a finite universe. Journal of the American Statistical Association 47: 663-685. [Crossref] [Google Scholar]
14. Rosenbaum PR, Rubin DB (1983) The central role of the propensity score in observational studies for causal effects. Biometrika 70: 41-55. [Crossref] [Google Scholar]

15. Robins JM, Rotnitzky A, Zhao LP (1994) Estimation of regression coefficients when some regressors are not always observed. Journal of the American Statistical Association 89: 846-866. [Crossref] [Google Scholar]
16. Hernán MA, Hernández-Díaz S, Robins JM (2004) A structural approach to selection bias. Epidemiology 15: 615-625. [Crossref] [Google Scholar]
17. Cole SR, Hernán MA (2008) Constructing inverse probability weights for marginal structural models. American Journal of Epidemiology 168: 656-664. [Crossref] [Google Scholar]
18. Robins JM, Hernán MA, Brumback B (2000) Marginal structural models and causal inference in epidemiology. Epidemiology 11: 550-560.[Google Scholar] [Indexed]
19. Lin DY, Lee-Jen W (1989) The robust inference for the Cox proportional hazards model. Journal of the American Statistical Association 84: 1074-1078. [Crossref] [Google Scholar]
20. Binder DA (1992) Fitting Cox's proportional hazards models from survey data. Biometrika 79: 139-147. [Crossref] [Google Scholar]
21. Shu D, Young JG, Toh S, Wang R (2021) Variance estimation in inverse probability weighted Cox models. Biometrics 77: 1101-1117. [Crossref] [Google Scholar] [Indexed]
22. Liu L, Nevo D, Nishihara R, Cao Y, Song M, et al. (2018) Utility of inverse probability weighting in molecular pathological epidemiology. European Journal of Epidemiology 33: 381-392. [Crossref] [Google Scholar] [Indexed]
23. Lunn M, McNeil D (1995) Applying Cox regression to competing risks. Biometrics 51: 524-532. [Crossref] [Google Scholar] [Indexed]
24. Wang M, Kuchiba A, Ogino S (2015) A meta-regression method for studying etiological heterogeneity across disease subtypes classified by multiple biomarkers. American Journal of Epidemiology 182: 263-270. [Crossref] [Google Scholar] [Indexed]
25. Chen RJ, Ding T, Lu MY, Williamson DFK, Jaume G, et al. (2024) Towards a general-purpose foundation model for computational pathology. Nature Medicine 30: 850-862. [Crossref] [Google Scholar] [Indexed]
26. Filiot A, Ghermi R, Olivier A, Jacob P, Fidon L, et al. (2023) Scaling self-supervised learning for histopathology with masked image modeling. [Crossref] [Google Scholar]
27. Xu H, Usuyama N, Bagga J, Zhang S, Rao R, et al. (2024) A whole-slide foundation model for digital pathology from real-world data. Nature 630: 181-188. [Crossref] [Google Scholar] [Indexed]
28. Lu MY, Chen B, Williamson DFK, Chen RJ, Liang I, et al. (2024) A visual-language foundation model for computational pathology. Nature Medicine 30: 863-874. [Crossref] [Google Scholar] [Indexed]
29. Wang X, Yang S, Zhang J, Wang M, Zhang J, et al. (2022) Transformer-based unsupervised contrastive learning for histopathological image classification. Medical Image Analysis 81: 102559. [Crossref] [Google Scholar] [Indexed]
30. Zhou J, Wei C, Wang H, Shen W, Xie C, et al. (2022) iBOT: Image BERT pre-training with online tokenizer. International Conference on Learning Representations. [Crossref] [Google Scholar]
31. Lu MY, Williamson DFK, Chen TY, Chen RJ, Barbieri M, Mahmood F (2021) Data-efficient and weakly supervised computational pathology on whole-slide images. Nature Biomedical Engineering 5: 555-570. [Crossref] [Google Scholar] [Indexed]
32. Ilse M, Tomczak JM, Welling M (2018) Attention-based deep multiple instance learning. 35th International Conference on Machine Learning 80: 2127-2136. [Google Scholar]
33. Campanella G, Hanna MG, Geneslaw L, Miraflor A, Silva VWK, et al. (2019) Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. Nature Medicine 25: 1301-1309. [Crossref] [Google Scholar] [Indexed]
34. Kather JN, Pearson AT, Halama N, Jäger D, Krause J, et al. (2019) Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. Nature Medicine 25: 1054-1056. [Crossref] [Google Scholar] [Indexed]
35. Sirinukunwattana K, Domingo E, Richman SD, Redmond KL, Blake A, et al. (2021) Image-based Consensus Molecular Subtype (imCMS) classification of colorectal cancer using deep learning. Gut 70: 544-554. [Crossref] [Google Scholar] [Indexed]
36. Pei-Chen T, Tsung-Hua L, Kun-Chi K, Fang-Yi S, Tsung-Lu ML, et al. (2023) Histopathology images predict multi-omics aberrations and prognoses in colorectal cancer patients. Nature Communications 14: 2102. [Crossref] [Google Scholar] [Indexed]
37. Chen RJ, Lu MY, Wang J, Williamson DFK, Rodig SJ, et al. (2022) Pathomic fusion: An integrated framework for fusing histopathology and genomic features for cancer diagnosis and prognosis. IEEE Transactions on Medical Imaging 41: 757-770. [Crossref] [Google Scholar] [Indexed]
38. Chen RJ, Lu My, Williamson DFK, Chen TY, Lipkova J, et al. (2022) Pan-cancer integrative histology-genomic analysis *via* multimodal deep learning. Cancer Cell 40: 865-878. [Crossref] [Google Scholar] [Indexed]
39. Moertel CG, Fleming TR, Macdonald JS, Haller DG, Laurie JA, et al. (1990) Levamisole and fluorouracil for adjuvant therapy of resected colon carcinoma. New England Journal of Medicine 322: 352-358. [Crossref] [Google Scholar] [Indexed]

40. Drysdale E (2022) SurvSet: An open-source time-to-event dataset repository. [Crossref] [Google Scholar] [Indexed]
41. Harrell FE, Califf RM, Pryor DB, Lee KL, Rosati RA (1982) Evaluating the yield of medical tests. JAMA 247: 2543-2546. [Crossref] [Google Scholar] [Indexed]
42. The Cancer Genome Atlas Network (2012) Comprehensive molecular characterization of human colon and rectal cancer. Nature 487: 330-337. [Crossref] [Google Scholar] [Indexed]
43. Uno H, Cai T, Pencina MJ, D'Agostino RB, Wei LJ (2011) On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data. Statistics in Medicine 30: 1105-1117. [Crossref] [Google Scholar] [Indexed]
44. Weisenberger DJ, Siegmund KD, Campan M, Young J, Long TI, et al. (2006) CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. Nature Genetics 38: 787-793. [Crossref] [Google Scholar] [Indexed]
45. Guinney J, Dienstmann R, Wang X, Reyniès A, Schlicker A, et al. (2015) The consensus molecular subtypes of colorectal cancer. Nature Medicine 21: 1350-1356. [Crossref] [Google Scholar] [Indexed]
46. Graf E, Schmoor C, Sauerbrei W, Schumacher M (1999) Assessment and comparison of prognostic classification schemes for survival data. Statistics in Medicine 18: 2529-2545. [Crossref] [Google Scholar] [Indexed]
47. Vickers AJ, Elkin EB (2006) Decision curve analysis: A novel method for evaluating prediction models. Medical Decision Making 26: 565-574. [Crossref] [Google Scholar] [Indexed]
48. Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, et al. (2024) TRIPOD+AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ 385: e078378. [Crossref] [Google Scholar] [Indexed]