Survival estimate models of cancer patients with spinal metastases: big data analytics to aid daily clinical practice (abridged secondary publication)

KYH Kwan *, CL Cheung, KMC Cheung, R Luo, DMC Poon, TW Lam, TC Lam

KEY MESSAGE

We used machine learning to predict 6-month survival after radiotherapy in 10 537 patients with spinal metastases. Our model demonstrated 88.5% accuracy in distinguishing survival outcomes. Its promising performance could enhance decision making for radiotherapy.

Hong Kong Med J 2025;31(Suppl 3):S12-3
HMRF project number: 07181576

[1] KYH Kwan, [2] CL Cheung, [1] KMC Cheung, [3] R Luo, [4] DMC Poon, [3] TW Lam, [5] TC Lam

[1] *Department of Orthopaedics and Traumatology, School of Clinical Medicine, The University of Hong Kong, Hong Kong SAR, China*
[2] *Department of Pharmacology and Pharmacy, The University of Hong Kong, Hong Kong SAR, China*
[3] *Department of Computer Science, The University of Hong Kong, Hong Kong SAR, China*
[4] *Department of Clinical Oncology, The Chinese University of Hong Kong, Hong Kong SAR, China*
[5] *Department of Clinical Oncology, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China*

* Principal applicant and corresponding author: kyhkwan@hku.hk

Introduction

Spinal metastases affect 60% to 70% of patients with advanced cancer.[1] Malignant spinal cord compression is associated with high morbidity and mortality.[2] To mitigate the adverse outcomes of spinal metastases, intensification of radiotherapy and surgery is proposed in addition to chemotherapy. Radiotherapy is utilised to control pain and improve outcomes in spinal metastasis.[3] Decision making regarding surgery for spinal metastases is complex. Treatment intensification is associated with substantial morbidity and mortality, particularly for spinal surgery.[4,5] Surgeries in an emergency setting are resource-intensive, requiring specialised surgeons and prolonged postoperative rehabilitation.[6] Similarly, intensification of radiotherapy using stereotactic body radiotherapy enables excellent local tumour control.[7] However, it requires increased resources in imaging and dosimetry, which are scarce in public hospital settings. Therefore, an accurate survival estimate is essential for the spinal care team to determine the most appropriate treatment for this heterogeneous patient group. For patients with short survival of ≤3 months,[8] invasive treatment with major decompressive surgery near the end of life is likely to cause major morbidity and prolonged hospitalisation. Conventional palliative radiotherapy alone, delivered in 10 fractions, is sufficient for pain relief and stabilisation in 70% of patients.[9] For patients with longer estimated survival, treatment intensification is deemed necessary to mitigate severe complications of spinal metastases, which can lead to irreversible lower limb paralysis.[10]

This study aimed to construct an accurate survival prediction model based on a cohort of approximately 10 000 patients with spinal metastases in Hong Kong, who exhibit heterogeneous backgrounds and outcomes.

TABLE 1. Model performance across various evaluation metrics.

Metrics	438 features	Top 50 features	Top 40 features	Top 30 features	Top 20 features	Top 15 features	Top10 features
Area under the receiver operating characteristic curve	0.885	0.879	0.879	0.879	0.871	0.864	0.839
Log-loss	0.421	0.433	0.438	0.436	0.447	0.464	0.495
Root mean square error	0.1353	0.140	0.141	0.141	0.144	0.148	0.162
Precision-recall	0.804	0.801	0.796	0.796	0.795	0.790	0.761
F1 Score	0.825	0.822	0.818	0.818	0.817	0.813	0.788

Methods

We constructed a prediction model using the Light Gradient Boosting Machine (a Microsoft-developed gradient boosting framework) for structured data and the bag-of-words model for unstructured data. The Light Gradient Boosting Machine is recognised for its advanced network communication, efficient memory utilisation, rapid training process, and effective overfitting prevention measures, facilitated by a histogram-based algorithm, leaf-wise growth strategy, gradient-based one-side sampling, and exclusive feature bundling. These mechanisms provide efficient solutions for high-dimensional datasets. For text data, we used the bag-of-words model to transform raw text into term frequency matrices, using a 10-fold cross-validation approach for model evaluation. Model optimisation was achieved through hyperparameter tuning using grid search and refinement of feature extraction from unstructured text. This process involved creating a 'stop words' list and incorporating n-grams of varying lengths to capture more meaningful information. The model was designed to maximise prediction accuracy while maintaining robustness and generalisability.

Results

After matching each patient with a unique reference key and excluding those with missing data, we obtained 10 537 unique reference keys. Each patient was treated as a unit of observation in the prediction problem: patient data collected prior to the diagnosis date were used to construct the independent variables, whereas

TABLE 2. The 10 most relevant features for survival after radiotherapy in our model.

Feature	Split	Gain	Quantile	Types	Explanation
Albumin	284.2	9.87	11.0-55.0 g/L	Numerical	Total amount of albumin in the blood (last test result before radiotherapy or most recent result)
lnratio_ab	237.6	7.63	0-12	Numerical	Neutrophil-to-lymphocyte ratio (absolute lymphocyte and absolute neutrophil counts from the last test result before radiotherapy or most recent result)
Codes_v66_7	168.4	7.43	0-35 times	Numerical	Number of inpatient diagnosis/prognosis codes (ICD-9-CM Diagnosis Code V66.7: encounter for palliative care)
Sodium	233.6	3.79	108.12-159 mmol/L	Numerical	Amount of sodium in the blood (last test result before radiotherapy or most recent result)
Codes_486	124.4	3.21	0-16 times	Numerical	Number of inpatient diagnosis/prognosis codes (ICD-9-CM Diagnosis Code 48.6: other resection of the rectum)
Duration_between_ip_and_RT	205	2.93	0-688 days	Numerical	Number of days between the discharge date of the latest hospitalisation before radiotherapy and the radiotherapy date
Codes_99_25	202.8	2.10	0-72 times	Numerical	Number of inpatient diagnosis/prognosis codes (ICD-9-CM Diagnosis Code 99.25: injection or infusion of a cancer chemotherapeutic substance)
Alkaline phosphatase	291	2.09	17-5971 U/L	Numerical	Amount of alkaline phosphatase in the blood. This enzyme is predominantly found in the liver, bones, kidneys, and digestive system. When the liver is damaged, alkaline phosphatase may leak into the bloodstream (last test result before radiotherapy or most recent result).
Codes_198_5	179.2	2.04	0-99 times	Numerical	Number of inpatient diagnosis/prognosis codes (ICD-9-CM Diagnosis Code 198.5: secondary malignant neoplasm of bone and bone marrow)
Phosphate	314	2.03	0.19-3.94 U/L	Numerical	A phosphate test measures the amount of phosphate in your blood (the last test result before radiotherapy or the latest one).

survival status 6 months after assessment served as the dependent variable.

The mean age of the 10 537 patients with spinal metastases was 64±13 (range, 5-101) years. Using one-hot encoding, independent variables were split into 3248 features. The importance of these features was ranked based on the amount of information gained. The model was then trained using the 438 features for which information gain was >0. The area under the receiver operating characteristic curve was the primary metric for assessing model performance in classification. The overall score of the model was 0.885. Other metrics including log-loss, root mean square error, precision-recall, and F1 score served as auxiliary measurement metrics. The performance metrics for this model, as well as those for models using only the top 50, 40, 30, 20, 15, and 10 features, respectively, are shown in Table 1. The most important feature was the total amount of albumin in the blood from the last test result before radiotherapy (or the most recent result). Discussions with practitioners confirmed that the top 10 features were most relevant for survival after radiotherapy (Table 2).

Conclusion

We used machine learning to predict 6-month survival rates after radiotherapy in patients with spinal metastases, based on a dataset of 10 537 patients. Our model demonstrated 88.5% accuracy in distinguishing survival outcomes. Its promising performance could enhance decision making for radiotherapy.

Funding

This study was supported by the Health and Medical Research Fund, Health Bureau, Hong Kong SAR Government (#07181576). The full report is available from the Health and Medical Research Fund website (https://rfs2.healthbureau.gov.hk).

References

1. Sutcliffe P, Connock M, Shyangdan D, Court R, Kandala N, Clarke A. A systematic review of evidence on malignant spinal metastases: natural history and technologies for identifying patients at high risk of vertebral fracture and spinal cord compression. Health Technol Assess 2013;17:1-274.
2. George R, Jeba J, Ramkumar G, Chacko AG, Tharyan P. Interventions for the treatment of metastatic extradural spinal cord compression in adults. Cochrane Database Syst Rev 2015;2015:CD006716.
3. Lam TC, Uno H, Krishnan M, et al. Adverse outcomes after palliative radiation therapy for uncomplicated spine metastases: role of spinal instability and single-fraction radiation therapy. Int J Radiat Oncol Biol Phys 2015;93:373-81.
4. Kim JM, Losina E, Bono CM, et al. Clinical outcome of metastatic spinal cord compression treated with surgical excision ± radiation versus radiation therapy alone: a systematic review of literature. Spine (Phila Pa 1976) 2012;37:78-84.
5. Quraishi NA, Rajagopal TS, Manoharan SR, Elsayed S, Edwards KL, Boszczyk BM. Effect of timing of surgery on neurological outcome and survival in metastatic spinal cord compression, Eur Spine J 2013;22:1383-8.
6. Fehlings MG, Nater A, Holmer H. Cost-effectiveness of surgery in the management of metastatic epidural spinal cord compression: a systematic review. Spine (Phila Pa 1976) 2014;39(22 Suppl 1):S99-S105.
7. Haley ML, Gerszten PC, Heron DE, Chang YF, Atteberry DS, Burton SA. Efficacy and cost-effectiveness analysis of external beam and stereotactic body radiation therapy in the treatment of spine metastases: a matched-pair analysis. J Neurosurg Spine 2011;14:537-42.
8. Fehlings MG, Nater A, Tetreault L, et al. Survival and clinical outcomes in surgically treated patients with metastatic epidural spinal cord compression: results of the prospective multicenter AOSpine study. J Clin Oncol 2016;34:268-76.
9. Hoskin P, Misra V, Hopkins K, et al. SCORAD III: randomized noninferiority phase III trial of single-dose radiotherapy (RT) compared to multifraction RT in patients (pts) with metastatic spinal canal compression (SCC). J Clin Oncol 2017;35(Suppl):LBA10004.
10. Malviya A, Gerrand C. Evidence for orthopaedic surgery in the treatment of metastatic bone disease of the extremities: a review article. Palliat Med 2012;26:788-96.