A comprehensive study about the prognosis predicting scoring systems of the vertebral metastasis patients, and the development of a novel calculator

PhD theses

Tamás Mezei M.D.

Semmelweis University
János Szentágothai Doctoral School of Neurosciences

Supervisor: Péter Banczerowski, M.D., D.Sc.
Co-supervisor: Péter Pollner, Ph.D.

Official reviewers:
Ildikó Vastagh, M.D., Ph.D.
Róbert Veres, M.D., Ph.D.

Head of the Complex Examination Committee:
Dániel Bereczki, M.D., D.Sc.

Members of the Complex Examination Committee:
Csaba Ertsey, M.D., Ph.D.
Attila György Bagó, M.D., Ph.D.

Budapest
2022

1 Introduction

Due to the increasing knowledge of cancer diseases, the study of their physiological and pathophysiological processes, the emergence of new biological therapies, the development of oncological surgery, the life expectancy of cancer patients has increased by 20-30% on average. However, this prolonged survival time leads to new problems and complications, which are presumably due to the accumulating tumour cell mutations, including the increasing metastatic potential. After the liver and lungs, the spinal column is the third most commonly affected part of the human body by metastases, which can affect nearly 70% (!) of cancer patients. 10% of vertebral metastases may produce clinical symptoms caused by metastatic epidural (spinal cord, root) compression which can result in a significant decrease in patients' quality of life, making them unsuitable for further oncological therapy and thus reducing life expectancy.

Surgical treatment of spinal metastases is palliative in nature. Our goal is to restore the stability of the spine, to preserve and/or improve neurological status by the decompression of the spinal cord and other nerve elements. However, choosing the personalized treatment method is not an easy task, we need to consider the potential complications, the advantages and disadvantages of the intervention, and the general and oncological condition of the patients.

Prognostic systems are widely used in the field of clinical medicine. In the case of cancer patients, a system that accurately estimates life expectancy has particular importance, as we can plan

therapy in a personalized way with knowledge of life prospects. Prognosis predicting scoring systems for patients with spinal metastases seek to participate as an objective factor in the decision-making process by estimating survival durations.

Several forms of risk calculators have emerged since the 1990's and have been an actively researched field in the clinical sciences of spinal surgery ever since. There are basic similarities in the function and composition of the different systems (e.g., each contains the type of primary tumour), the difference being the combination of factors, the formation of subgroups and the statistical methodology of system creation.

The "first generation" of prognosis predicting systems are based on weighting the survival values of clinical prognostic factors. Newer, "second generation" systems use predictive tools for weighting, which are developed by advanced statistical modelling as a result of machine learning. Only recent results are available for newer systems that require external validation.

In my doctoral thesis I dealt with the operation and clinical applicability of classically structured systems.

2 Objectives

1. Development of a single-centre database of surgically treated spinal metastatic patients with previously unpublished size in Europe to ensure better knowledge of demographic data and statistically relevant evaluability.

2. Identification of new, easily accessible and usable prognostic factors.

3. Study of the functionality and usability of the best-known and most-cited first-generation prognostic systems in a large Hungarian retrospective cohort with spinal metastases.

4. In our previous studies, we found that known systems are able to separate patients reproducibly according to their survival time, however, significant inaccuracies can be observed in their predictive ability. Our goal is to investigate and correct the cause of the inaccuracy.

5. Internationally, the performance of surgeries is subject to stricter indication factors than in our institute, where we also operate on patients with poor life expectancies, if we can achieve a potentially significant symptomatic improvement (reduction of pain, resolution of motor and/or autonomic disorders) with a lower surgical burden. Thus, our aim was to answer the question of how long it is worthwhile to

provide our patients with operative care and to what extent we can expect therapeutic benefits from the interventions.

6. The main goal of my doctoral dissertation was to develop a new, prognosis predicting and therapeutic decision supporting scoring system, as a summary of our research results and experiences, based on a reliable statistical methodology which is intended not only for surgeons but also for colleagues involved in the care of patients with spinal metastases (e.g., oxiologist, neurologist, radiotherapist, oncologist, anaesthesiologist).

3 Methods

A single-centre based, retrospective database was created at the National Institute of Mental Disorders, Neurology and Neurosurgery (Department of Neurosurgery of Semmelweis University) from adult patients with epidural vertebral metastases who underwent surgery.

The study has so far taken place in 2 periods. We started developing the database in 2016. We collected data from January 1, 2008 to December 31, 2015, on 337 patients and 384 surgical interventions with the use of our electronic patient registration system (Hospitaly®). After evaluating the materials and publishing our results, we started expanding the database in 2018, and finally we identified information about 454 patients and 521 surgical interventions between January 1, 2008 and January 1, 2018. A minimum follow-up period of one and a half years was set, so mortality data were requested from the Hungarian Cancer Registry in June 2019.

More features were collected from each patient, both for the pre-, intra, and postoperative periods, and 40 potential prognostic factors per patient were identified.

All patients were scored retrospectively according to 1990's Tokuhashi, 2005's Tokuhashi, Tomita, modified Bauer, van der Linden, Katagiri, and to the new scoring system proposed here.

3.1 Statistical analysis

Simple **descriptive statistical methods** were used to evaluate demographic data. **Fischer exact tests** were performed when we examined the correlation of discrete data (e.g., in the assessment of therapeutic benefit). **Binomial tests** were performed when ascertaining whether a frequency characteristic differed from the expected value. For survival studies, **Kaplan-Meier analysis** was performed, and the difference between the survival curves of each group was assessed by a **log-rank test**. A significance threshold of 5% was set during the studies.

As a first step in developing the scoring system, we used **Kaplan-Meier analysis** and **log-rank tests** to screen for factors that were unable to differentiate patients according to their survival probability. The large degrees of freedom of some parameters were reduced with the use of agglomerative clustering (e.g., numerical factors such as serum albumin levels were classified into two or three significantly different categories, and related component analysis was used to group the factors for discrete factor variables).

In the next step, we performed a **network science-based correlation analysis** of the prognostic factors which influenced the survival probability of the patients. The predictors of the proposed system were determined by the maximum click search method. This provided the greatest help in the selection of the factor combinations.

All parameters, which were not discarded, were combined into a **uni- and multivariate Cox hazard model**. Taking into account the interaction between the factors, it was possible to reduce the number

of significant parameters with a backward elimination procedure. The proportionality assumption of the Cox model was tested at a 15% significance level. For the most suitable combinations, the difference of the weighting parameters observed between the uni- and multivariate analysis was minimal.

We compared the predicting power of the novel scoring system with other systems with a **cut-off analysis**, calculating **Uno's C-index** at several different censoring times. A higher value at a given time point indicates that the scoring system remains sensitive at high specificity for assessment of survival until the end of the chosen time window. **Delta statistics** and the **Integrated Discrimination Index** were calculated to quantify the difference between the prediction power of the scoring systems.

All calculations were done with R software (version 3.4.4).

4 Results

4.1 Analysis of prediction fault of the revised Tokuhashi system

In our previous studies, we demonstrated that prognosis predicting systems for patients with spinal metastases are able to separate patients according to their survival times, however, their predictive power was still unsatisfactory. The best results were given by the revised Tokuhashi system, so considering it as a basic model, we focused on identifying the errors of the above system and correcting them.

The revised Tokuhashi system distinguishes 3 prognostic categories, which differ significantly in terms of survival probabilities ($p<0.001$), however, its predictive power was 60.5% on average. Detailed survival values for each group were examined separately. The most significant difference between the predicted and real survival parameters was observed in the "conservative category", which includes patients with poor life expectancies. The literature survival rate (OS=0.15 at 180 days) is below the lower 95% confidence interval of the real OS value obtained from the survival analysis of our population (OS=0.38, 95%CI=0.307-0.471), which means a significant difference in term of survival. The main difference in this category that the system only recommends conservative therapy (chemotherapy, radiotherapy), but in our case all patients underwent at least a minimally invasive decompression surgery. Our population passes the 0.15 OS value on day 475 (95%CI=359-796

days) instead of 180 days. This represents a significant increase in survival time of nearly 10 months, attributed to the potential benefits achieved by surgical treatments.

Since this level of bias was observed for the category, we aimed to modify the point limits of the categories to improve the predictive ability. Survival curves of patients assigned to scores in the conservative and palliative groups were examined to develop a new, changed upper limit of the conservative group and a new, lower limit of the palliative group. Based on our results, we recommend the changing of the scoring method of the revised Tokuhashi system as follows:

- Excisional category: 12-15 points (unchanged)
- **Palliative category: 7-11 points** (formerly 9-11)
- **Conservative category: 0-6 point(s)** (formerly 0-8)

4.2 Examination of functional outcome in the patient population with poor life expectancy

We can state that even in patients with poor life expectancy, a significant increase in survival time can be achieved with surgical intervention. Due to preserved or restored neurological functions, patients' quality of life is not reduced to such an extent that their oncology care is contraindicated, so they can continue or start the treatment.

In all 132 patients in the conservative category, pain occurred as a symptom of the disease, which was alleviated in 130 cases by surgery, indicating a significant ($p < 1e-5$) change by binomial test.

Despite significant improvements in quality of life (98-99%), the chances of motor dysfunction resolving are only around 50%.

Existing autonomic dysfunction (incontinence, retention) showed the worst results in terms of chances of improvement. Compared with pain (p=0.001, OR=5.01, CI95%=1.86–13.97), and limb weakness (p<1e–5, OR=10.64, CI95%=3.81–31,35), the latter two problems can be solved significantly more successfully.

4.3 Creating our scoring system

Based on the knowledge gained during our previous studies, after expanding our database, we developed our population-specific prognosis predicting and therapeutic option recommending scoring system.

In the case of 32 prognostic factors (out of 40) we confirmed the ability to influence survival by Kaplan-Meier analyses and log-rank tests. To weight the prediction ability, we performed a Cox analysis (with Schoenfeld plots to narrow the number of usable factors), and finally found the 5 most suitable factors with network-science based correlation analysis (maximum clique search technique). The results of the uni- and multivariate Cox analysis of these 5 factors differ very slightly, thus indicating their ability as an independent predictor.

The scoring methodology was developed using the **β-coefficient values** because of their additive properties obtained by the Cox analysis, so the given scores really represent a value correlated with survival. Our system and its usage are shown in **Table 1.**

1. Table The proposed risk calculator and our therapeutic option recommendation

Prognostic factors	**Risk calculator point(s)**
Primary tumour type	
1st category: hematologic malignancies, thyroid	**0**
2nd category: breast and prostate cancer	**6**
3rd category: angiosarcoma, chondrosarcoma, osteosarcoma, Ewing-sarcoma, bladder, colorectal, epipharynx, kidney, liver, melanoma, ovarium, uterus, parotis, stomach, unknown origin	**12**
4th category: cervix, oesophagus, larynx, lung, pancreas, penis, sino-nasal	**19**
Age at the time of the operation	
1st category: under 40 years	**0**
2nd category: over 40 years	**8**
Mobility	
1st category: able to walk	**0**
2nd category: bed bound (incapable of walking)	**5**
Presence of visceral metastasis	
1st category: no	**0**
2nd category: yes	**4**
Serum total protein level	
1st category: normal level (›6,5 g/dl)	**0**
2nd category: hypoproteinemia (‹6,5 g/dl)	**3**

2. Table continued The proposed risk calculator and our therapeutic option recommendation

Prognostic groups of the risk calculator	**Risk calculator total point(s)**
Good survival probability group	**0-17** Surgical treatment is recommended. Based on the radiologic findings - we should be as radical as we can ("excisional" surgery)
Moderate survival probability group	**18-24** Surgical treatment is recommended, "less-invasive" techniques are preferred in this category
Poor survival probability group	**25-39** Primarily, palliative onco-radiotherapy is recommended. Minimal invasive spine surgery techniques could be suggested if we can achieve a significant improvement in the quality of life (resolve pain syndrome, restore neurological function, etc.)
Spinal Instability Neoplastic Score (SINS) can help assess the need for stabilization.	

The survival predicting ability of our scoring system was compared with the well-known systems in the literature, namely the original Tokuhashi, revised Tokuhashi, Tomita, modified Bauer, van der Linden and Katagiri. Our methodology was Uno's C-statistic, which is an estimation of the ROC analysis, developed for "small" case numbers and censored data (still alive at the end of the study period).

We found that the new system has the highest C-index value over the whole time interval. It is the most sensitive and specific in terms of predictive ability having **0.706** value at 10 years (CI95%=0.679-0.733), followed by Katagiri **0.658** (CI95%=0.628-0.688), revised Tokuhashi **0.645** (CI95%=0.616-0.675), modified Bauer **0.635** (CI95%=0.605-0.664), Tomita **0.624** (CI95%=0.593– 0.654), original Tokuhashi **0.615** (CI95%=0.586-0.644) and van der Linden **0.567** (CI95%=0.538-0.596) systems. The difference between the predictive power of the investigated system was assessed using D-statistics, and IDI plots were generated. The new system turns out to be significantly stronger than the other 6 well-known systems. It is **5%** more accurate than Katagirin, **6%** more accurate than the revised Tokuhashi, **7%** more accurate than the modified Bauen, **8%** more accurate than the Tomita, **9%** more accurate than the original Tokuhashi and **14%** more accurate than the van der Linden system.

5 Conclusions

The treatment of patients with metastatic spinal tumours should be considered from a complex spinal surgical, oncological and radio therapeutical perspective. The reason for the importance of the disease is the increasing incidence and prevalence data, the progressive neurological symptoms caused by instability and compression of nerve elements, which lead to a significant deterioration of perfomance, making our patients unsuitable for further oncotherapy. Prognostic scoring systems can help in the therapeutic decision-making process.

In my doctoral thesis, I carried out a comprehensive study of the prognostic systems of patients with spinal metastases, the main findings, novelties and results of which are as follows:

1. We collected the prognostic factors of adult spinal metastases patient's and described the characteristics of the population in the framework of a one-centre based retrospective clinical study of a size not previously published in the European literature.

2. Among the factors significantly influencing survival, we identified new ones (e.g., serum protein level, serum sodium level, Charlson comorbidity index), and by re-clustering subgroups of existing factors, we achieved an increase in their estimation ability.

3. We found that several clinical factors are unsuitable for estimating the outcome of patients with spinal metastasis, out of which I would like to highlight the presence and number of extraspinal bone metastases and the affected level of the spine and the number of intraspinal metastases.

4. We examined the predictive power of the most cited prediction systems in the literature in a Hungarian spinal metastasis population. We found that the systems were able to differentiate patients according to their survival time, however, we observed significant predictive bias that limits their usefulness.

5. As a sample system, we examined the revised Tokuhashi system in detail. The main reason for the prediction inaccuracy is suspected to lie in the scoring methodology and treatment recommendations, as the system is too strict for the "poor life expectancy" patient group, for whom only supportive therapy was recommended. Based on our results, even patients with a poor prognosis should be treated surgically, as we can achieve a significant increase in life expectancy, with an improvement in quality of life (surgery can significantly reduce patient pain).

6. We proposed to modify the point limits of the prognostic categories of the revised Tokuhashi system in order to achieve better predictive power.

7. Based on the literature research and our own results, it can be stated that the development of population-specific systems is necessary in order to eliminate environmental factors, genetic and epigenetic features, and social differences.

8. **We have established and validated our prognosis predicting and therapy recommending system, for which we have also provided a webpage. https://emk.semmelweis.hu/gerincmet**

9. We presented a state-of-the-art statistical methodology for the development of a clinical prediction system that could provide an excellent basis for other large centres' own risk calculators.

6 Bibliography of the candidate's publications

6.1 Publications related to the dissertation

1. Czigléczki G, **Mezei T**, Pollner P, Horváth A, Banczerowski P
   Prognostic factors of surgical complications and overall survival of patients with metastatic spinal tumor
   World Neurosurgery, 113:e20-e28. (2018)
   doi: 10.1016/j.wneu.2018.01.092
   IF: 1,723

2. Pollner P*, Horváth A*, **Mezei T**, Banczerowski P, Czigléczki G
   Analysis of four scoring systems for the prognosis of patients with metastasis of the vertebral column
   World Neurosurgery, 112:e675-e682. (2018)
   doi: 10.1016/j.wneu.2018.01.124
   IF: 1,723

3. **Mezei T,** Horváth A, Pollner P, Czigléczki G, Banczerowski P
   Research on the predicting power of the revised Tokuhashi system: how much time can surgery give to patients with short life expectancy?
   International Journal of Clinical Oncology, 25 (4):755-764. (2020)
   doi: 10.1007/s10147-019-01612-w
   IF: 3,402

4. **Mezei T**, Báskay J, Pollner P, Horváth A, Nagy Z, Cziglóczki G, Banczerowski P; [New, innovative prognosis calculator for patients with metastatic spinal tumors].
   Ideggyógyászati Szemle, 75 (3-04):117-127. (2022)
   doi: 10.18071/isz.75.0117
   IF: 0,427

6.2 Publications on other topics

1. **Mezei T**, Hajdu M, Cziglóczki G, Lotz G, Kocsis J, Kulka J, Horváth A; Sterile, abscess-like cerebral lesion during trastuzumab therapy after HER2 status switch in a triple negative breast cancer patient: a case report and literature review
   BMC Cancer, 20 (1):615 (2020)
   doi: 10.1186/s12885-020-07114-7
   IF: 4,430

2. **Mezei T**, Mészáros D, Pollner P, Bagó A, Fedorcsák I, Banczerowski P, Sipos L; [Supplementary valproate therapy for glioma patients. An alternative opportunity to enhance the efficiency of radio-chemotherapy]
   Orvosi Hetilap, 2021; 162(24): 960–967. (2021)
   doi: 10.1556/650.2021.32110
   IF: 0,540

7 Funding

This research was supported by the Doctoral School of Semmelweis University. The project received partial funding from the "EFOP-3.6.3-VEKOP-16-2017-00009, Az orvos-, egészségtudományi- és gyógyszerészképzés tudományos műhelyeinek fejlesztése" scholarship.

The research was partially financed by the Thematic Excellence Programme (TKP 2020-4.1.1.-TKP2020) of the Ministry for Innovation and Technology in Hungary, within the framework of the Digital Biomarker thematic programme of the Semmelweis University.