



OPEN

## Serum protein triplet TGF- $\beta$ 1, TIMP-1, and YKL-40 serve as diagnostic and prognostic profile for astrocytoma

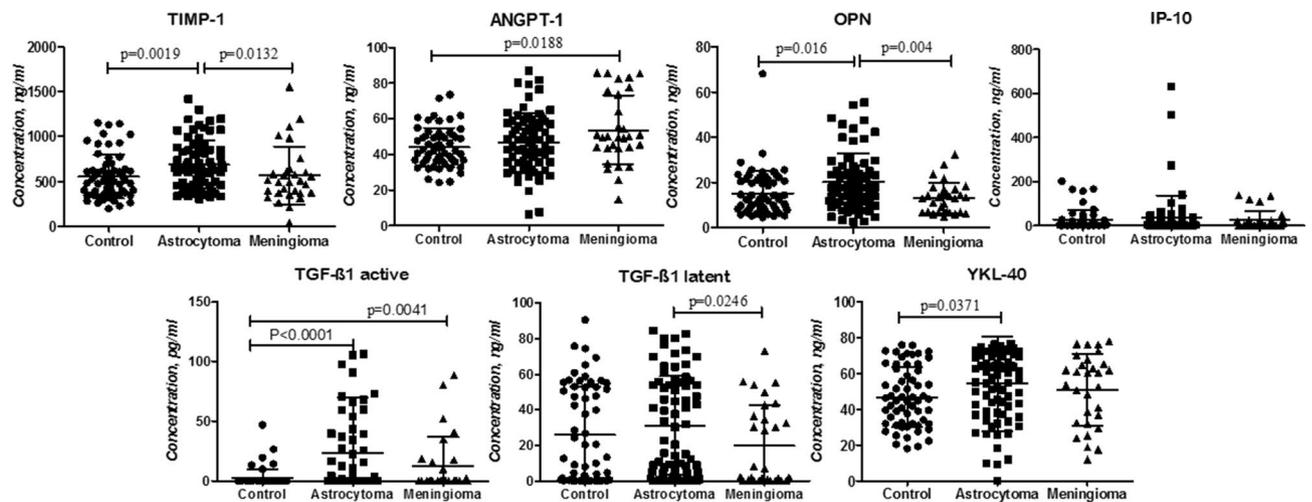
Rūta Urbanavičiūtė<sup>1✉</sup>, Rūta Zabitaitė<sup>1</sup>, Algimantas Kriščiukaitis<sup>2</sup>, Vytenis-Pranas Deltuva<sup>1</sup> & Daina Skiriūtė<sup>1</sup>

Astrocytoma is the most common glial tumour of the CNS. The most malignant form is grade IV Astrocytoma, also called Glioblastoma. Due to its heterogeneity, aggressiveness and lethal nature scientists are trying to find less invasive methods for early prediction of tumour onset, recurrence, response to therapy and patients' survival. Here, applying decision tree classification algorithm we performed astrocytoma specific protein profile analysis on serum proteins TIMP-1, active and latent form of TGF- $\beta$ 1, IP-10, ANGPT-1, OPN, and YKL-40 using enzyme-linked immunosorbent detection assay (ELISA). Results have demonstrated that astrocytoma specific profile consisted of three proteins—active form of TGF- $\beta$ 1, TIMP-1 and YKL-40 and was able to correctly classify 78.0% (103/132) of sample and 83.3% (60/72) of astrocytoma sample. Calculating decision tree algorithm associated with astrocytoma patient survival, prediction model reached an accuracy of 83.3% (60/72). All together these results indicate that glioma detection and prediction from patient serum using glioma associated proteins and applying mathematical classification tools could be achieved, and applying more comprehensive research further could be implemented in clinic.

Astrocytoma is the most common type of brain tumour originating in astrocytes, which are star-shaped brain cells located in the cerebrum<sup>1</sup>. According to the World Health Organization (WHO) classification, based on histological parameters, astrocytoma is divided into grade II (diffuse astrocytoma), grade III (anaplastic astrocytoma) and grade IV (glioblastoma (GBM))<sup>2</sup>. A higher grade astrocytoma represents a worse prognosis and is more aggressive. Patients with grade II astrocytoma have a median survival of ~7–8 years. While patients with grade III astrocytoma have a median survival of ~2–3 years, and patients with GBM of only 9–14 months<sup>3</sup>. Delayed diagnosis is a major contributing factor to lower patients' survival. Mostly, astrocytoma is diagnosed by computer tomography or magnetic resonance imaging scan. Although these methods are reliable and used for many years, they are expensive and usually used after the tumour has spread a lot<sup>4</sup>. Therefore, it is important to find new non-invasive methods that could detect tumour in early stages and prognosticate its progression, and response to therapy. One of the most promising non-invasive methods could be detection of certain molecules, i.e. biomarkers, such as proteins in the patient's blood, however as to our knowledge none is used to diagnose astrocytoma yet.

A recent study done by our group<sup>5</sup> revealed promising blood serum candidate protein markers of astrocytoma progression. We constructed decision tree classification algorithm for the indication of presence of astrocytoma—specific serum protein profiles, and found that tissue inhibitor of matrix metalloproteinase 1 (TIMP-1), active transforming growth factor beta 1 (TGF- $\beta$ 1), C-X-C motif chemokine ligand 10 also known as interferon gamma-induced protein 10 (IP-10), angiopoietin-1 (ANGPT-1) and osteopontin (OPN) can be potential mediators of astrocytoma progression. In addition, several studies have confirmed that these proteins could be potential astrocytoma diagnostic biomarkers due to their biological functions during oncogenesis<sup>6–11</sup>. Besides, recently numerous studies emphasized chitinase-3-like protein 1 (CHI3L1), also known as YKL-40 as one of the most promising astrocytoma diagnostic and prognostic biomarker<sup>12–14</sup>. To strengthen the proposed model, we decided to include YKL-40. Additionally, model was supplemented with analytes from brain tumour

<sup>1</sup>Laboratory of Molecular Neurooncology, Neuroscience Institute, Lithuanian University of Health Sciences, Eiveniu str. 4, 50161 Kaunas, Lithuania. <sup>2</sup>Laboratory of Biophysics and Bioinformatics, Neuroscience Institute, Lithuanian University of Health Sciences, Eiveniu str. 4, 50161 Kaunas, Lithuania. ✉email: ruta.urbanaviciute@ismuni.lt



**Figure 1.** TIMP-1, ANGPT-1, OPN, IP-10, TGF-β1 active, TGF-β1 latent and YKL-40 proteins expression levels in astrocytoma (grade II-IV) patients; healthy control and benign meningioma patients' peripheral blood serum. Middle line in each group represents mean, lower and upper lines—standard deviation (SD).

meningioma in order to check model specificity to glioma. Since the majority of proteins are implicated in many different biological processes it is important to investigate sorted out protein's cancer-type specificity (i. e. glioma) comparing to cancers having different type of pathogenesis mechanism. Meningioma is dural-based tumour that arise from arachnoid cap or meningeothelial cells<sup>15</sup> and shares the same location as glioma, while yet having distinct nature, meningioma could serve as convenient model to study molecular markers specificity to glioma.

Here we performed seven serum proteins—TIMP-1, active and latent forms of TGF-β1, IP-10, ANGPT-1, OPN, and YKL-40—analysis in astrocytic glioma patient sample applying conventional ELISA assay. The aim of this study was the construction of the diagnostic and prognostic glioma pathology associated serum protein profiles applying decision tree method.

## Results

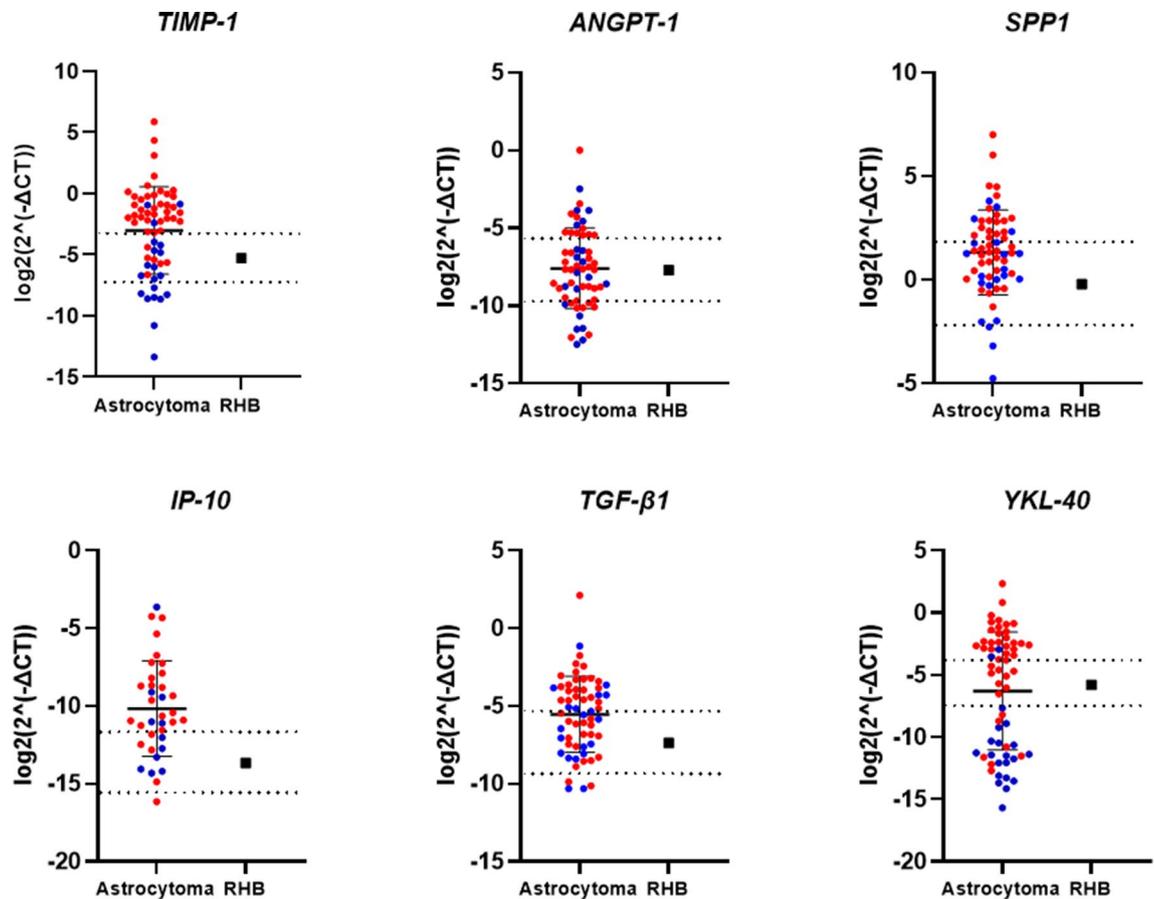
**Protein levels in different brain tumour entities.** Protein expression levels in astrocytoma and meningioma patient blood serum using ELISA method were investigated. The expression levels among different brain tumours (astrocytoma (grade II-IV) vs meningioma (grade I)) and control group are demonstrated in Fig. 1. TIMP-1, OPN, YKL-40, and active form of TGF-β1 protein levels in astrocytoma serum substantially differed from levels in healthy control group. TIMP-1, OPN, and latent form of TGF-β1 levels in astrocytoma differed from meningioma group. ANGPT-1 as well as active form of TGF-β1 demonstrated difference between healthy control and meningioma group.

**mRNA expression of protein coding genes in tumour tissue.** To elucidate if the protein level in blood serum is related with the presence of brain tumours, we investigated the mRNA expression of *TIMP-1*, *ANGPT1*, *SPP1* (OPN coding gene), *IP-10*, *TGF-β1* and *TIMP-1* in astrocytoma samples as well as in healthy human brain tissue.

In total mRNA expression of *TIMP-1*, *ANGPT1*, *SPP1*, *TGF-β1* and *YKL-40* genes was analysed in 61 patients with different grade astrocytoma: 20 grade II diffuse astrocytoma, and 41 grade IV astrocytoma (glioblastoma), and in case of *IP-10*—in 11 grade II diffuse astrocytoma, and 24 glioblastoma samples. All genes mRNA level was quantified in tumour tissue using qRT-PCR, and compared with mRNA level in Human normal brain RNA sample “FirstChoice Human Brain Reference RNA” (Ambion), which was a pool of RNAs assembled from 12 donors from several brain regions, as described by the manufacturer. qRT-PCR showed stronger expression of *SPP1*, and *TGF-β1*, in most astrocytoma samples of both grade II and IV compared to normal brain sample. Elevated expression only in GBM samples compared to normal brain was visible in *TIMP-1*, *IP-10*, and *YKL-40* gene groups. *YKL-40* also shown downregulated expression in grade II astrocytoma tumour samples compared with normal brain tissue (Fig. 2). These data are in the line with one of protein expression in serum described above.

**Combined Analysis of Candidate Protein Levels in Astrocytoma Serum.** The decision tree classification analysis was performed using “Classification Learner” application provided by MathWorks MatLab on TIMP-1, ANGPT-1, OPN, IP-10, TGF-β1 active, TGF-β1 latent and YKL-40 protein expression levels in astrocytoma patients and healthy control groups. The results demonstrated classification tree composed of three proteins—active TGF-β1, TIMP-1, YKL-40, and correctly classified 78.0% (103/132) of the whole sample, 71.7% (43/60) of healthy control sample and 83.3% (60/72) of astrocytoma sample (Fig. 3).

Decision tree models were also applied using different tumour entities as variables, however, none of the models achieved higher accuracy than the previous one. Model of healthy group vs. patients with any type of brain cancer demonstrated 72.8% accuracy, healthy control vs. grade II astrocytoma vs. GBM—68.9%; healthy control vs. astrocytoma (grade II-IV) vs. meningioma—63.0%; healthy control vs. benign brain tumours (grade



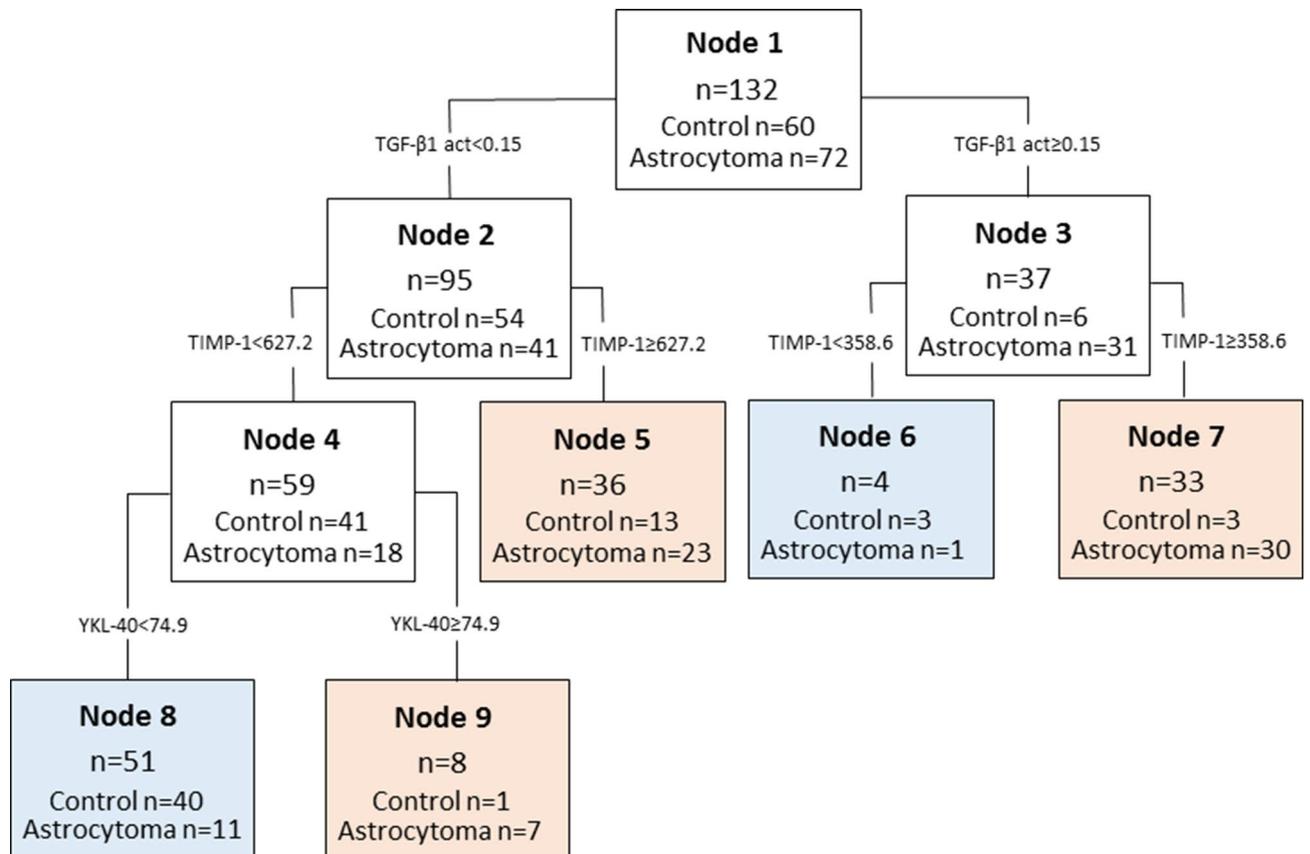
**Figure 2.** mRNA level of *TIMP-1*, *ANGPT1*, *SPPI*, *IP-10*, *TGF-β1* and *TIMP1* genes in astrocytoma and healthy human brain specimens. RHB—reference human brain RNA sample (pool of RNA from 12 healthy individuals). Blue dots in astrocytoma group indicates grade II astrocytoma samples, and red dots—samples of GBM. The upper and lower dotted lines indicate twofold higher and twofold lower mRNA expression points from RHB, respectively. Middle line in astrocytoma group represents mean, lower and upper lines—standard deviation (SD).

I meningioma and grade II astrocytoma) vs. GBM—56.8%, healthy control vs. GBM—75.8%; healthy control vs. grade II astrocytoma vs. meningioma—76.5%.

**Decision tree model specificity to brain tumours.** As earlier mentioned, meningioma in this study serves as a model to investigate the specificity of glioma molecular markers. Herein, we tested the algorithm's calculated for astrocytoma detection from blood serum and composed of active TGF-β1, TIMP-1 and YKL-40, ability to differentiate meningioma. Half of meningioma patients (15/30) were assigned to healthy control group, while another half—to astrocytoma group. Altogether, decision tree model for healthy control vs. astrocytoma (grade II-IV) vs. meningioma reached 63.0% accuracy with only 13.3% (4/30) of correct samples assigned to meningioma group, indicates model selectivity to glioma type cancer.

**Survival analysis.** Kaplan–Meier survival analysis demonstrated that astrocytoma patients with higher than median YKL-40 and latent form of TGF-β1 levels in blood serum survived significantly longer ( $p=0.031$  and  $p=0.035$ , respectively), while the lower than median protein concentration in patient serum indicated shorter patients' survival. TIMP-1, ANGPT-1 and active form of TGF-β1 demonstrated the tendency of patient survival association with the lower protein level in astrocytoma, but the difference was not significant. IP-10 and OPN haven't shown any associations (Fig. 4). Survival analysis in separate GBM group for studied serum proteins did not showed any differences (see Supplementary Fig. S1).

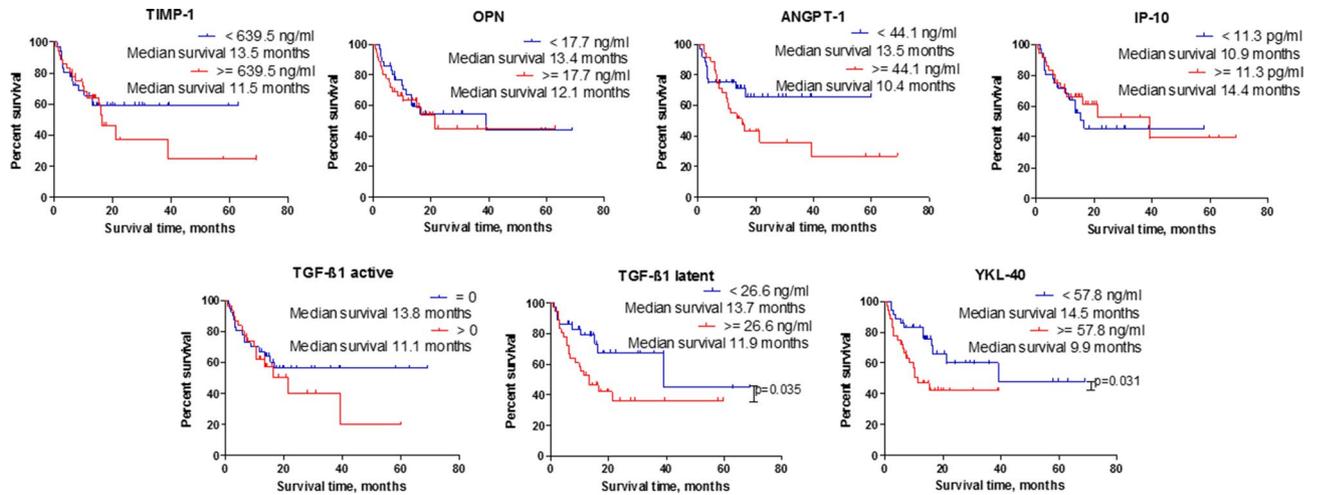
In the combined survival analysis, total ( $n=72$ ) astrocytoma sample was divided into two groups—patients who survived less than a year after the surgery ( $n=34$ ), and those who survived longer than a year ( $n=38$ ). Calculated decision tree algorithm with the highest accuracy of 83.3% (60/72) consisted of YKL-40, latent and active TGF-β1 and TIMP-1 and predicted the survival of more than one year with a specificity of 94.7% (36/38), while shorter than a year survival—with a specificity of 70.6% (24/34) (Fig. 5.).



**Figure 3.** Diagnostic astrocytoma decision tree model, composed of active (act) TGF- $\beta$ 1, TIMP-1 and YKL-40 proteins in astrocytoma and healthy control serum. Tree was calculated from the results of 132 participants (60—controls and 72—astrocytomas). The number of all participants assigned to the specific class is represented by n, below—the number of participants in healthy control and astrocytoma group. Rose colour nodes represent classes, which predict participant having astrocytoma, blue nodes—healthy participant.

## Discussion

In decades of glioma research none blood serum or plasma marker for disease detection or prediction have been introduced to clinic, on the other hand, many studies have tried to find single marker or complexes, which could be specific enough to recognize glioma tumour from blood. In this study the potential of blood serum proteins for detection and prediction of astrocytoma was tested. Promising candidate proteins—TIMP-1, ANGPT-1, OPN, IP-10, active form of TGF- $\beta$ 1, latent form of TGF- $\beta$ 1 and YKL-40—were chosen according to the earlier our laboratory findings<sup>5</sup> and broad literature review<sup>6–14,16–19</sup>. Here we showed that protein levels of TIMP-1, OPN, YKL-40, and active form of TGF- $\beta$ 1 in astrocytoma patient serum substantially differed as compared to control (healthy) group, while the differences between healthy control and meningioma group were observed only for ANGPT-1 and TGF- $\beta$ 1 active. Altogether these results reiterate numerous other studies. Tissue inhibitor of metalloproteinases (TIMPs) together with matrix metalloproteinases (MMPs) regulates degradation of extracellular matrix (ECM) and are very important factors in tumour invasion processes. Lin and colleagues showed that TIMP-1 level was increased in glioma patient's plasma as compared to normal controls<sup>19</sup>. Next protein in our investigation is OPN well known as a modulator of apoptosis, extracellular matrix degradation, cell migration and neovascularization<sup>20</sup>. Elevated levels of OPN in patient blood have been reported in many cancer types<sup>21</sup> as well as in astrocytoma<sup>16</sup>. YKL-40 has been considered as one of the most promising proteins in astrocytoma detection. It is 40-kDa secreted glycoprotein, which is produced by cancer cells and is involved in many cancer related processes like cell proliferation, apoptosis, differentiation, angiogenesis, and regulation of extracellular tissue remodelling<sup>12,13</sup>. The differences in YKL-40 protein serum level between healthy subjects and glioma patients have been demonstrated already in 2002 by Tanwar group<sup>13</sup>, and a comprehensive analysis of Zhi-Qiang Li<sup>14</sup> group confirmed YKL-40 protein being a valuable biomarker for glioma patient prognosis. Next protein in our analysis, namely TGF- $\beta$ 1 in blood serum can be measured in two forms—latent and active. Together with Latency Associated Peptide (LAP) the mature TGF- $\beta$ 1 dimer form a latent complex, in which TGF- $\beta$ 1 is unable to bind its receptors. In a few cell types, latent TGF- $\beta$ 1 can be activated, when release from the LAP occurs. Activation of the cytokine implies release of mature TGF- $\beta$ 1 from LAP<sup>22</sup>. Dysregulated TGF- $\beta$  signalling was reported in many patients, and is possibly related with initiation and progression of many cancers including glioma<sup>9</sup>. As to our knowledge, none of these proteins are reported to associate with meningioma as possible serum molecular marker, however abnormal levels of TIMP-1 and OPN have been reported in meningioma tumour tissue<sup>23,24</sup>.



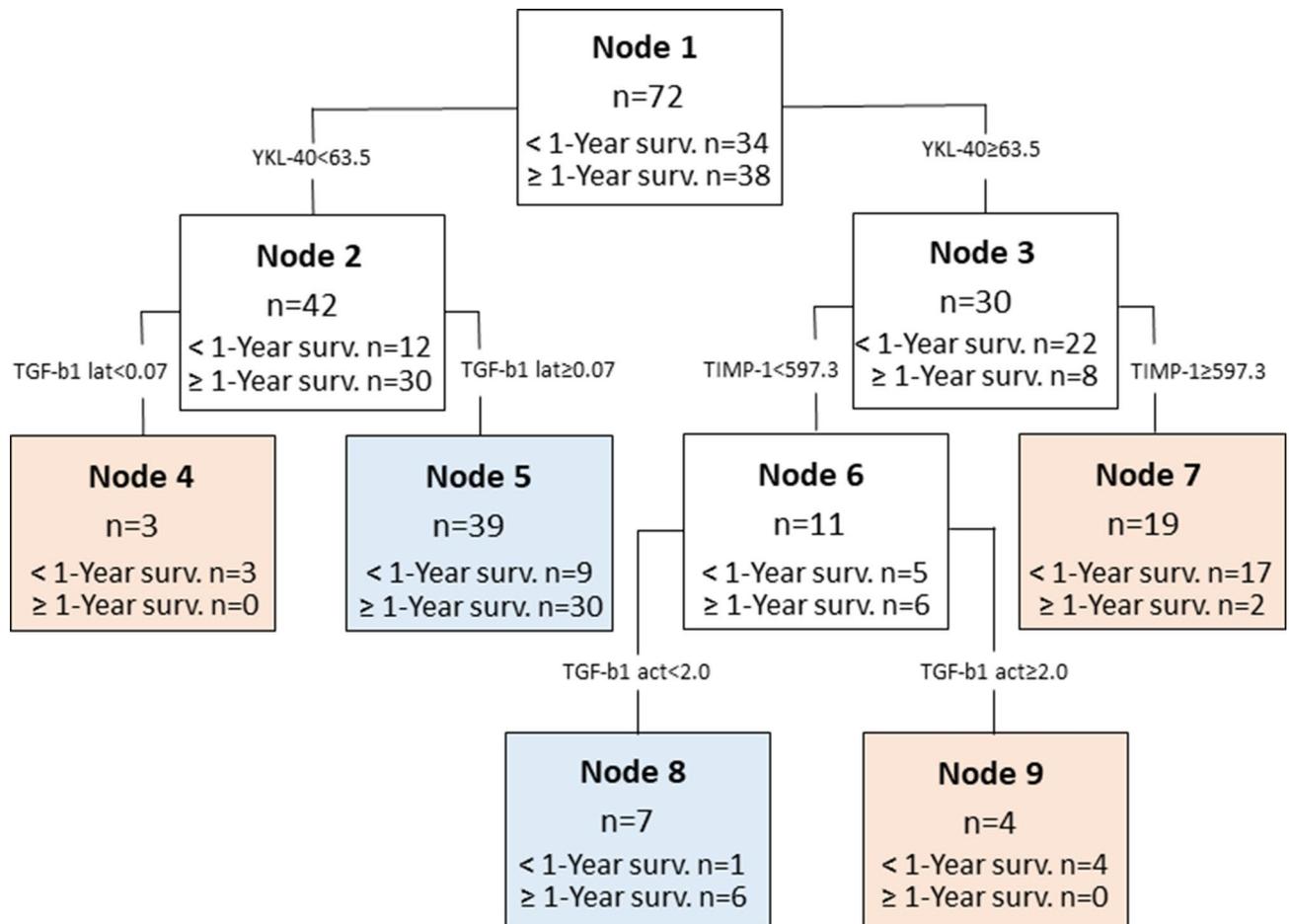
**Figure 4.** Survival dependency on TIMP-1, OPN, ANGPT-1, IP-10, active and latent TGF- $\beta$ 1, and YKL-40 proteins concentration in astrocytoma patients peripheral blood serum. TIMP-1 ( $n = 72$ ) low ( $< 639.5$  ng/mL) versus high ( $\geq 639.5$  ng/mL) protein level patient group with median survival of 13.5 months versus 11.5 months, respectively; log-rank test;  $\chi^2 = 0.535$ ;  $df = 1$ ;  $p = 0.464$ ; OPN ( $n = 72$ ), low ( $< 17.7$  ng/mL) versus high ( $\geq 17.7$  ng/mL) protein level patient group with median survival of 13.4 months versus 12.1 months, respectively; log-rank test;  $\chi^2 = 0.231$ ;  $df = 1$ ;  $p = 0.631$ ; ANGPT-1 ( $n = 72$ ), low ( $< 44.1$  ng/mL) versus high ( $\geq 44.1$  ng/mL) protein level patient group with median survival of 13.5 months versus 10.4 months, respectively; log-rank test;  $\chi^2 = 2.679$ ;  $df = 1$ ;  $p = 0.102$ ; IP-10 ( $n = 72$ ), low ( $< 11.3$  pg/mL) vs. high ( $\geq 11.3$  pg/mL) protein level patient group with median survival of 10.9 months versus 14.4 months, respectively; log-rank test;  $\chi^2 = 0.293$ ;  $df = 1$ ;  $p = 0.589$ ; active TGF- $\beta$ 1 ( $n = 72$ ), low (0 pg/mL) versus high ( $> 0$  pg/mL) protein level patient group with median survival of 13.8 months versus 11.1 months, respectively; log-rank test;  $\chi^2 = 0.617$ ;  $df = 1$ ;  $p = 0.432$ ; latent TGF- $\beta$ 1 ( $n = 72$ ), low ( $< 26.6$  ng/mL) versus high ( $\geq 26.6$  ng/mL) protein level patient group with median survival of 13.7 months versus 11.9 months, log-rank test;  $\chi^2 = 4.468$ ;  $df = 1$ ;  $p = 0.035$ ; YKL-40 ( $n = 72$ ), low ( $< 57.8$  ng/mL) versus high ( $\geq 57.8$  ng/mL) protein level patient group with median survival of 14.5 months versus 9.9 months, respectively; log-rank test;  $\chi^2 = 4.672$ ;  $df = 1$ ;  $p = 0.031$ ; Blue colour—low protein level, red colour—high protein level.

Herein, as most promising protein profile for both low and high grade astrocytoma detection was composed of TGF- $\beta$ 1 active, TIMP-1 and YKL-40. Decision tree algorithm correctly classified 78.0% (103/132) of the sample, 71.7% (43/60) of healthy controls sample and 83.3% (60/72) of astrocytoma sample. Interestingly, fibrous meningioma sample following this model were equally assigned to both healthy control and astrocytoma groups, indicating probable protein profile specificity to glioma cancer, rather than brain cancer in common.

Earlier we demonstrated<sup>5</sup> decision tree which consisted of ANGPT-1, TIMP-1, IP-10 and active TGF- $\beta$ 1 proteins, algorithm reached accuracy of 73.5% (75/102) and correctly classified 79.7% (47/59) of all astrocytomas and 65.1% (28/43) of all healthy controls. Accuracy of the decision tree calculated using expression levels of those four proteins in the independent sample set demonstrated even greater results: the accuracy of 75.8%, with 100 of 132 correctly assigned participants (healthy vs. astrocytoma patients) to the right group (see Supplementary Fig. S2). Furthermore, the model demonstrated glioma type cancer specificity as 53.3% (16/30) of fibrous meningioma samples were assigned to control group, while 47.7% (14/30)—to astrocytoma group using the same decision tree parameters. Elstner with colleagues proposed promising serum protein profile, consisting of BMP2, HSP70, and CXCL10 (IP-10), which was able to distinguish GBM group with the specificity and sensitivity of 89% and 96% from control one, respectively<sup>25</sup>. Pérez-Larraya in 2014 published 2-step diagnostic algorithm using glial fibrillary acidic protein (GFAP) and YKL-40 which differentiated patients with GBM from those with nonglial brain tumours with a sensitivity of 65% and a specificity of 78%<sup>26</sup>.

Despite many innovations in the modern medicine and combined treatment applying surgery, chemotherapy and radiotherapy, brain gliomas remain lethal, aggressively growing tumours leading to short patient survival. The possibility almost noninvasively prognosticate glioma patient survival could help to optimize tumour treatment decisions. Here we presented decision tree model which was able to correctly predict 60 out of 72 astrocytoma patients (83.3%) one year survival with the accuracy of 94.7% (36/38). Furthermore, prediction tree consisted of the same proteins as the astrocytoma detection algorithm that is TGF- $\beta$ 1 active, TIMP-1, YKL-40, with the addition of the protein of latent form TGF- $\beta$ 1. Several promising protein profiles prognosticating patient survival were also demonstrated earlier. The group of Lin have shown that the panel of four cytokines—IL-7, IL1R4/ST2, sgp130 and MCP-1 substantially correlated with GBM patient's overall survival<sup>27</sup>, as well as Elstner and his group proposed the protein profile formed of TSP1, HSP70, and IGFBP3, which was associated with more than 15 months survival after GBM tumour resection<sup>25</sup>.

All together these results indicate that glioma detection and prediction from patient serum using glioma associated proteins and applying mathematical classification tools could be achieved and further comprehensive research could help to implement molecular prediction tools into clinical usage.



**Figure 5.** Potential prognostic profile of serum proteins YKL-40, latent (lat) and active (act) forms of TGF- $\beta$ 1 and TIMP-1 for astrocytoma patients. The overall accuracy of the algorithm reached 83.3% (60/72), less than one year survival accuracy reached 70.6% (24/34), while more than 1 year survival accuracy reached 94.7% (36/38). The number of participants assigned to the specific class is represented by n. Rose colour represents classes, predicting shorter than a year patient survival, blue—longer than a year patient survival group.

## Materials and methods

**Study group and blood collection.** Study group was collected from 2014 to 2019 at the Neurosurgery department of the Hospital of Lithuanian University of Health Sciences Kaunas Clinics with the informed consent of all participants. Investigation methodologies were approved by Kaunas Regional Biomedical Research Ethics Committee (No. P2-9/2003) and were carried out under the Declaration of Helsinki. In the study group of 162 serum samples 72 were histopathologically confirmed astrocytoma (12 of which—grade II astrocytoma; 60 of which—grade IV astrocytoma), 30 samples fibrous meningioma (benign brain tumour) and 60 control samples from healthy volunteers. Median/mean age of healthy control group was 51.7/51.8 yr., median/mean age of astrocytoma group—62.9/58.1 yr. and median/mean age of meningioma group was 61.2/62.9 yr. Blood from both astrocytoma and meningioma patients were collected before tumour resection and any kind of treatment, while after operation the standard treatment was applied (radiotherapy plus chemotherapy). For control group healthy volunteer individuals with no evidence of brain or any other type of cancer along with no indications of inflammation or infection were chosen. Collected blood serum samples were allowed to clot from 30 min. to 1 h at RT and then were centrifuged at 1300xg for 10 min. Serum was collected, transferred to fresh Eppendorf tubes and stored at  $-80^{\circ}\text{C}$  until further investigation stage.

**Evaluation of serum protein levels.** For protein concentration measurements in participant's blood serum R&D Systems ELISA kits were used: for TIMP-1—DY970-05 kit, for ANGPT-1—DY923 kit, for OPN—DY1433 kit, for YKL-40—DY2599 kit, both forms of TGF- $\beta$ 1—DY240-05 kits and for IP-10—DY266-05 kit. For each protein detection, different dilution factor was used according to the manufacturers' recommendations, other groups studies, or pilot experiments<sup>28–33</sup>. Serum for TIMP-1 was diluted 400 folds; ANGPT-1—20 folds; OPN—20 folds; YKL-40—50 folds; TGF- $\beta$ 1 active—undiluted; TGF- $\beta$ 1 latent—undiluted; IP-10—undiluted. To avoid multiple freeze thaw cycles before the first experiment each serum sample was aliquoted according to chosen dilution factor for every investigated protein, including two additional tubes with serum quantity equal to one undiluted sample in case of unsuccessful experiment and stored at  $-80^{\circ}\text{C}$  until further investigation stage. Every sample was measured in duplicate. Protein concentration of TGF- $\beta$ 1 in serum was measured for both

active and latent forms. All experimental steps and data analysis were performed according to the manufacturers' recommendations<sup>28–33</sup>. Optical densities were measured at a wavelength of 450 nm using a “Sunrise” absorbance reader (“Tecan Trading”, Switzerland).

**RNA extraction, cDNA synthesis and qRT-PCR.** RNA was extracted from 100 mg of nitrogen frozen tumor tissue using “mirVana miRNA Isolation Kit” (Invitrogen, USA). cDNA was synthesized from 2 µg RNA using “High-Capacity cDNA Reverse Transcription Kit” (Applied Biosystems, USA) with Ribolock RNase inhibitor. Reverse transcription was performed with hexamer primers and MultiScribe Reverse Transcriptase according to manufacturer's instructions. PCR applying Sybr-Green I (“Maxima SYBR Green/ROX qPCR Master Mix”, Thermo Scientific, Lithuania) assay was carried out in 12 µl of total reaction volume including 15 ng of cDNA. Primers were designed using “PerlPrimer3” and synthesized by Metabion International (Germany). Primer sequences could be obtained under request. qRT-PCR was performed on AB7500 fast platform (Applied Biosystems). Gene expression level was calculated applying dCt method and geometric mean of ACTB and GAPDH Ct values was used for data normalization.

**Statistical analysis.** To analyse gathered data, software packages GraphPad Prism version 6.0 (San Diego, CA, USA) and SPSS Statistics 19 (Chicago, IL, USA) were used. First, distribution of data sets were verified using Kolmogorov–Smirnov test. To evaluate differences of each investigated protein levels in blood serum between two independent groups, both with normally distributed data parametric Student's t-test was applied. To compare groups, where one or both data sets are abnormal the non-parametric Mann–Whitney test was used. Kaplan–Meier method was used to estimate astrocytoma patient survival function. All analysed protein groups were divided into lower (< median value) and higher (≥ median value) protein expression groups, according to the median values of each protein level and the log-rank test was used to compare the survival curves. Survival time was calculated starting from the date of operation until death, or the date of project closure.

For Decision Tree analysis was applied “Classification Learner” application provided by MathWorks as part of Machine Learning Toolbox in MATLAB<sup>34</sup>.

### Data availability

The datasets generated during and analysed during the current study are available from the corresponding author on a request.

Received: 21 January 2021; Accepted: 8 June 2021

Published online: 23 June 2021

### References

- Liu, Y. *et al.* Multidimensional analysis of gene expression reveals TGFβ1-induced EMT contributes to malignant progression of astrocytomas. *Oncotarget* **5**, 12593–12606 (2014).
- Louis, D. N. *et al.* The 2016 world health organization classification of tumors of the central nervous system: a summary. *Acta Neuropathol.* **131**, 803–820 (2016).
- Ren, T., Lin, S., Wang, Z. & Shang, A. Differential proteomics analysis of low- and high-grade of astrocytoma using iTRAQ quantification. *Onco. Targets. Ther.* **9**, 5883–5895 (2016).
- Davis, M. E. Glioblastoma: overview of disease and treatment. *Clin. J. Oncol. Nurs.* **20**, 1–8 (2016).
- Vaitkiene, P. *et al.* Identification of astrocytoma blood serum protein profile. *Cells* **9**, 16 (2019).
- Groft, L. L. *et al.* Differential expression and localization of TIMP-1 and TIMP-4 in human gliomas. *Br. J. Cancer* **85**, 55–63 (2001).
- Aaberg-Jessen, C. *et al.* Low expression of tissue inhibitor of metalloproteinases-1 (TIMP-1) in glioblastoma predicts longer patient survival. *J. Neurooncol.* **95**, 117–128 (2009).
- Crocker, M. *et al.* Serum angiogenic profile of patients with glioblastoma identifies distinct tumor subtypes and shows that TIMP-1 is a prognostic factor. *Neuro. Oncol.* **13**, 99–108 (2011).
- Han, J., Alvarez-Breckenridge, C. A., Wang, Q. E. & Yu, J. TGF-β signaling and its targeting for glioma treatment. *Am. J. Cancer Res.* **5**, 945–955 (2015).
- Liu, Z., Kuang, W., Zhou, Q. & Zhang, Y. TGF-β1 secreted by M2 phenotype macrophages enhances the stemness and migration of glioma cells via the SMAD2/3 signalling pathway. *Int. J. Mol. Med.* **42**, 3395–3403 (2018).
- Enderlin, M. *et al.* TNF-α and the IFN-γ-inducible protein 10 (IP-10/CXCL-10) delivered by parvoviral vectors act in synergy to induce antitumor effects in mouse glioblastoma. *Cancer Gene Ther.* **16**, 149–160 (2009).
- Johansen, J. S., Schultz, N. A. & Jensen, B. V. Plasma YKL-40: a potential new cancer biomarker?. *Future Oncol.* **5**, 1065–1082 (2009).
- Tanwar, M. K., Gilbert, M. R. & Holland, E. C. Gene expression microarray analysis reveals YKL-40 to be a potential serum marker for malignant character in human glioma. *Cancer Res.* **62**, 4364–4368 (2002).
- Zhao, Y.-H. *et al.* YKL-40 in high-grade glioma: prognostic value of protein versus mRNA expression. *Glioma* **1**, 104 (2018).
- Shaikh, N., Dixit, K. & Raizer, J. Recent advances in managing/understanding meningioma [version 1; referees: 2 approved]. *F1000Research* vol. 7 (2018).
- Sreekanthreddy, P. *et al.* Identification of potential serum biomarkers of glioblastoma: serum osteopontin levels correlate with poor prognosis. *Cancer Epidemiol. Biomarkers Prev.* **19**, 1409–1422 (2010).
- Reiss, Y., Machein, M. R. & Plate, K. H. The role of angiopoietins during angiogenesis in gliomas. *Brain Pathol.* **15**, 311–317 (2005).
- Schneider, T., Sailer, M., Ansoorge, S., Firsching, R. & Reinhold, D. Increased concentrations of transforming growth factor beta1 and beta2 in the plasma of patients with glioblastoma. *J. Neurooncol.* **79**, 61–65 (2006).
- Lin, Y. *et al.* Plasma levels of tissue inhibitor of matrix metalloproteinase-1 correlate with diagnosis and prognosis of glioma patients. *Chin. Med. J. (Engl)* **126**, 4295–4300 (2013).
- Wai, P. Y. & Kuo, P. C. Osteopontin: Regulation in tumor metastasis. *Cancer Metastasis Rev.* **27**, 103–118 (2008).
- Elevated plasma osteopontin in metastatic breast cancer associated with increased tumor burden and decreased survival—PubMed. <https://pubmed.ncbi.nlm.nih.gov/9815727/>.
- Stockis, J., Dedobbeleer, O. & Lucas, S. Role of GARP in the activation of latent TGF-β1. *Mol. Biosyst.* **13**, 1925–1935 (2017).
- Arikök, A. *et al.* Osteopontin expressions correlate with WHO grades and predict recurrence in meningiomas. *Springer* <https://doi.org/10.1007/s10014-013-0152-2> (2013).

24. Kachra, Z. *et al.* Expression of matrix metalloproteinases and their inhibitors in human brain tumors. *Clin. Exp. Metastasis* **17**, 555–566 (1999).
25. Elstner, A. *et al.* Identification of diagnostic serum protein profiles of glioblastoma patients. *J. Neurooncol.* **102**, 71–80 (2011).
26. Gállego Pérez-Larraya, J. *et al.* Diagnostic and prognostic value of preoperative combined GFAP, IGFBP-2, and YKL-40 plasma levels in patients with glioblastoma. *Cancer* **120**, 3972–3980 (2014).
27. Lin, Y. *et al.* A panel of four cytokines predicts the prognosis of patients with malignant gliomas. *J. Neurooncol.* **114**, 199–208 (2013).
28. Human TIMP-1 DuoSet ELISA DY970–05: R&D Systems. [https://www.rndsystems.com/products/human-timp-1-duoset-elisa\\_dy970](https://www.rndsystems.com/products/human-timp-1-duoset-elisa_dy970).
29. Human Osteopontin (OPN) DuoSet ELISA DY1433: R&D Systems. [https://www.rndsystems.com/products/human-osteopontin-opn-duoset-elisa\\_dy1433](https://www.rndsystems.com/products/human-osteopontin-opn-duoset-elisa_dy1433).
30. Human Chitinase 3-like 1 DuoSet ELISA DY2599: R&D Systems. [https://www.rndsystems.com/products/human-chitinase-3-like-1-duoset-elisa\\_dy2599](https://www.rndsystems.com/products/human-chitinase-3-like-1-duoset-elisa_dy2599).
31. Human CXCL10/IP-10 DuoSet ELISA DY266-05: R&D Systems. [https://www.rndsystems.com/products/human-cxcl10-ip-10-duoset-elisa\\_dy266](https://www.rndsystems.com/products/human-cxcl10-ip-10-duoset-elisa_dy266).
32. Human Angiopoietin-1 DuoSet ELISA DY923: R&D Systems. [https://www.rndsystems.com/products/human-angiopoietin-1-duoset-elisa\\_dy923](https://www.rndsystems.com/products/human-angiopoietin-1-duoset-elisa_dy923).
33. Human TGF-beta 1 DuoSet ELISA DY240–05: R&D Systems. [https://www.rndsystems.com/products/human-tgf-beta-1-duoset-elisa\\_dy240](https://www.rndsystems.com/products/human-tgf-beta-1-duoset-elisa_dy240).
34. Train models to classify data using supervised machine learning—MATLAB—MathWorks Switzerland. <https://ch.mathworks.com/help/stats/classificationlearner-app.html>.

## Acknowledgements

We appreciate Jūratė Žeglienė and all the nurses from the LUHS Hospital Kaunas Clinics Department of Neurosurgery for the kind assistance in collecting blood serum samples and patient data.

## Author contributions

R.U., and D.S. conception of the study and funding acquisition; R.Z. protein experimental part, ELISA assays; V.D. patient sample and data gathering; R.U., and A.K. data analysis and interpretation; R.U. article draft; D.S. writing and editing. All authors reviewed the manuscript.

## Funding

This research was funded by a grant (No. MIP-052/2015) from the Research Council of Lithuania and LUHS Science Fund for doctoral student project.

## Competing interests

The authors declare no competing interests. The corresponding author is responsible for submitting a competing interest statement on behalf of all authors of the paper.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-92328-3>.

**Correspondence** and requests for materials should be addressed to R.U.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021