Received:
4 March 2017

Accepted:
8 May 2017

<https://doi.org/10.1259/bjr.20170168>

Cite this article as:

Liu Z, Zhang J. Age, choline-to-*N*-acetyl aspartate, and lipids-lactate-to-creatine ratios assemble a significant Cox's proportional-hazards regression model for survival prediction in patients with high-grade gliomas. *Br J Radiol* 2017; **90**: 20170168.

LETTER TO THE EDITOR

Age, choline-to-*N*-acetyl aspartate, and lipids-lactate-to-creatine ratios assemble a significant Cox's proportional-hazards regression model for survival prediction in patients with high-grade gliomas

ZHENYIN LIU, MD and JING ZHANG, MD

Department of Medical Imaging, Guangzhou Women and Children's Medical Center, Guangzhou Medical University, Guangzhou, China

Address correspondence to: Dr Jing Zhang

E-mail: dorzhangjing@163.com

(The Editors do not hold themselves responsible for opinions expressed by correspondents).

To the Editor,

With great interest, we read the article¹ *Age, choline-to-*N*-acetyl aspartate, and lipids-lactate-to-creatine ratios assemble a significant Cox's proportional-hazards regression model for survival prediction in patients with high-grade gliomas* (by Roldan-Valadez et al¹ 2016). In this work, the authors developed a Cox's models to predict survival of patients with high-grade gliomas.

In this article, Cox's model¹ ($N = 28$) predictors included 20 clinical and MR variables. When fitting a prediction model, it is well known that we pick up part of the idiosyncratic characteristics of the data along with the systematic

relationship between dependent and explanatory variables. This phenomenon is known as overfitting and generally occurs when a model is excessively complex relative to the amount of data available.² This is commonly referred to as the high-dimensional, low sample size (*i.e.* $p \gg n$) setting.³ Overfitting, which is characterized by high accuracy for a classifier when evaluated on the training data set but by low accuracy when assessed on a separate test set, has been recognized as a problem in $p \gg n$ settings.³ Overfitted models make the results of a research project appear better than they are, thus (falsely) impressing colleagues and referees.⁴ As a result, the authors are recommended to use variable selection to represent more precise fit of their raw data. Further study with a large cohort is required.

REFERENCES

1. Roldan-Valadez E, Rios C, Motola-Kuba D, Matus-Santos J, Villa AR, Moreno-Jimenez S. Choline-to-*N*-acetyl aspartate and lipids-lactate-to-creatine ratios together with age assemble a significant Cox's proportional-hazards regression model for prediction of survival in high-grade gliomas. *Br J Radiol* 2016; **89**: 20150502. doi: <https://doi.org/10.1259/bjr.20150502>
2. Bilger M, Manning WG. Measuring overfitting in nonlinear models: a new method and an application to health expenditures. *Health Econ* 2015; **24**: 75–85. doi: <https://doi.org/10.1002/hec.3003>
3. Subramanian J, Simon R. Overfitting in prediction models—is it a problem only in high dimensions? *Contemp Clin Trials* 2013; **36**: 636–41. doi: <https://doi.org/10.1016/j.cct.2013.06.011>
4. Chen S, McMullan G, Faruqi AR, Murshudov GN, Short JM, Scheres SH, et al. High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryomicroscopy. *Ultramicroscopy* 2013; **135**: 24–35. doi: <https://doi.org/10.1016/j.ultramic.2013.06.004>