Received: 17 January 2023 | Revised: 17 April 2023 | Accepted: 24 July 2023

DOI: 10.1002/cac2.12475

CANCER COMMUNICATIONS Open Access

LETTER TO THE EDITOR

Genomic information of children with malignant brain tumors for the prediction of length of hospitalization

Dear Editor:

Central nervous system tumors in the brain or spine are the most common solid tumors in children, which accounts for about 25% of cancers in children younger than 15 years of age, and are the most common cause of cancer deaths in children [1]. The 5-year survival rate for central nervous system neoplasms has increased dramatically to 74% for patients under 18 years old (97% for benign/borderline malignant tumors) in 2022, compared to 20% in the 1970s [1]. Without cure treatments and specific medications for many brain cancers, the dramatic increase of survival rate is largely due to improved hospital care and availability of clinical resources. The length of hospitalization of pediatric patients with brain cancer is an important indicator of prognosis as it reflects the required medical effort needed to care for these patients. The length of hospitalization is also critical for the healthcare system as hospital admission is a part of the care trajectory of the respective patients, related to the cost of medical care [2, 3]. According to the recent data released by the Association of American Medical Colleges (AAMC) in 2020/2021, an estimated physician shortage in the United States will be between 37,800 and 124,000 by 2034, including shortfalls in both primary and specialty care [4]. Global crisis of physician shortage has become a challenge for hospital management that cannot be ignored, which emphasizes the importance of the length of hospitalization prediction because it is a significant factor for hospital resource utilization.

Due to difficulties with biopsy, researchers have been making efforts in recent decades to identify predictive biomarkers for the disease prognosis and the length of hospitalization by minimal or non-invasive methods, such as blood biomarkers. Although some promising candidates have been reported, including panels of proteins and single biomarkers such as the genes glial fibrillary acidic protein (*GFAP*), interleukin-10 (*IL10*) and individual microRNAs (miRNAs), no robust and widely used blood-based biomarkers for malignant brain tumors were identified so far for clinical applications [5]. Due to the availability of advanced genomic technologies, genomic information is attracting extensive research attention to assist early and accurate diagnosis, clinical decision for treatment procedures, prognosis prediction, and efficient application of hospital resources. In this study, we conducted a genomic study on genome-wide DNA variants and length of hospitalization related to malignant brain tumor, leveraging one of the most powerful pediatric oncology resources assembled to date, i.e., whole-genome sequencing (WGS) of blood DNA from 1,243 children in the Gabriella Miller Kids First project (https://commonfund.nih.gov/kidsfirst). The goal of this study was to investigate the feasibility of predicting the length of hospitalization for cancer care by genomic information.

In this study, pediatric patients with birth defects (BD) were recruited by the Center for Applied Genomics (CAG) at the Children's Hospital of Philadelphia (CHOP) (Supplementary Methods). A total of 12 patients with both malignant brain tumors and leukemia/other malignant blood diseases were excluded to remove the noise from chromosome aberrations of cancer cells before the analysis. For further analysis, we included 157 BD probands who were diagnosed with at least one type of malignant brain tumors and at least one type of BD, while 767 patients with BD but cancer-free (BD-only) were used as the benign disease control group besides 319 healthy control subjects. Clinical information, such as age at diagnosis, gender/ethnicity ratios, and tumor locations, are described in Supplementary Figures S1-S2 and Supplementary Tables S1-S2. As disease comorbidity is commonly seen in children with both brain tumor and BD, we used BD-only samples as background level to reduce false positive signals unrelated to brain tumor in children without BD. The length of hospitalization in the brain tumor subjects was determined based on detailed information from

Abbreviations: AAMC, Association of American Medical Colleges; BD, birth defects; CAG, Center for Applied Genomics; EMR, Electronic medical records; GFAP, glial fibrillary acidic protein; IL10, interleukin 10; LncRNA, long non-coding RNAs; miRNAs, micro-RNA; RFE, recursive feature elimination; WGS, whole genome sequencing.

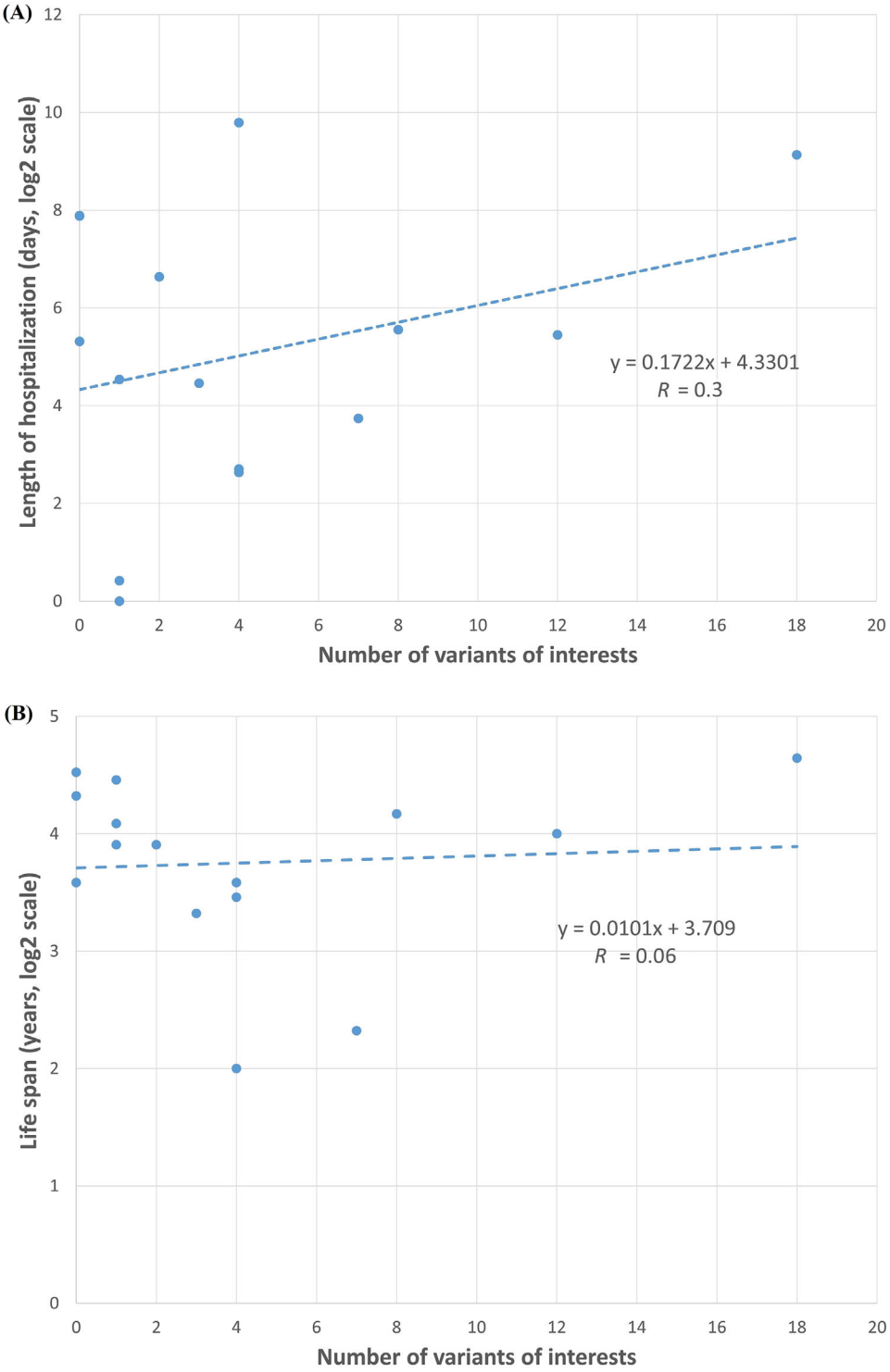

FIGURE 1 Correlation coefficient plots between the load of variants of interest and brain tumor-related length of hospitalization (A) or life (B) spans of 15 deceased patients with malignant brain tumors. The variants of interest are the same in Figure 1A and 1B.

electronic medical records (EMR) for each of the Center for Applied Genomics (CAG) recruited patients, as described in the Supplementary Methods, to ensure the length of hospitalization was attributed to the malignant brain tumor, but not due to other medical conditions. WGS was done at 30 × coverage for all the individuals using the genomics platform at the Broad Institute of Massachusetts Institute of Technology and Harvard University (Cambridge, MA, USA). The annotations for the variants were generated using the ANNOVAR software

developed by our group [6]. Recurrent variants exclusively occurred in the patients with malignant brain tumor were identified. The genomic variants were filtered based on their signal strengths associated with length of hospitalization using recursive feature elimination (RFE) model with random forest regressor as the estimator. A total of 1,760 variants were selected by the RFE algorithm selection processes and mapped to 695 genes/non-coding RNAs (Supplementary Table S3).

From the recurrent variants identified in children with brain tumors, markers related to the development and progression of cancers, including chloride channel accessory 4 (*CLCA4*), mucin 21, cell surface associated (*MUC21*), perforin 1 (*PRF1*), Galactosylceramidase (*GALC*), and fucosyltransferase 3 (*FUT3*), were identified (Supplementary Table S4). Meanwhile, consistent functional enrichment was identified from the Wnt signaling pathway of cancer, which was suggested by multiple modules with significantly adjusted *P* value, using the RNA-seq/proteome profiling co-expression networks for glioblastoma multiforme and glioma data from The Cancer Genome Atlas Program and The Clinical Proteomic Tumor Analysis Consortium [7]. The Wnt signaling regulates key cellular events during the development of the central nervous system [8] and is a key signaling pathway in glioblastoma targeted by multiple clinical trial drugs [9]. Another interesting finding was that long non-coding RNAs (lncRNAs) account for a significant portion of the variants of interest and corresponding genes. Among the 1,760 variants of interest, 404 (23%) fall into lncRNA loci, corresponding to 116 lncRNAs. These results suggested that non-coding RNAs, especially lncRNAs, might play more significant roles in determining the prognosis of malignant brain tumor than previously reported. A recent study has shown abnormal expression of lncRNAs in brain tumors, and lncRNA-related molecular mechanisms may guide brain tumor therapy [10].

Using an independent dataset for validation is essential to evaluate the reliability of the prediction model. For this purpose, we selected 15 deceased patients with malignant brain tumor who were genetically unrelated to the 157 children with brain cancer (Supplementary Table S5). Variants of interests identified in these 15 deceased patients are described in Supplementary Table S6. The life span was counted from the year of birth to the deceased year. Load of variants of interest showed a positive correlation to brain tumor-related length of hospitalization in the 15 patients (correlation coefficient = 0.3), indicating that with the more variants identified in our model, the longer time the patients were likely to be hospitalized due to their brain tumor-related medical conditions. Therefore, the reported variants of interest, especially those in genomic coding regions, could serve as predictor variables for the length of hospitalization. Our results demonstrated that the load of variants of interest was associated with brain tumor-related length of hospitalization (Figure 1A). However, no correlation was found between the load of variants of interest and life span (correlation coefficient = 0.06, Figure 1B). The lifespan is influenced by a complex interplay of genetic and clinical factors. Therefore, predicting it using genetic variants may be more challenging compared to predicting the length of hospitalization.

In summary, this study disclosed the correlation between genome wide variants and the length of hospitalization related to malignant brain tumor in children. We leveraged one of the largest pediatric oncology resources assembled to date. We demonstrated that the load of genetic variants of interest could predict the length of hospitalization of these patients. The predictive capability was supported by an independent validation dataset. The variants of interest identified in this study could serve as predictor variables for the length of hospitalization and to inform clinical decision making. In addition, our results highlighted the importance of lncRNAs related to prognosis of brain tumors. Additional data, including in-depth phenotypic information, may further improve the current prediction model with genomic information alone and will gain more power for this unique research resource to capture genetic variants with biological and clinical relevance.

DECLARATIONS

AUTHOR CONTRIBUTIONS

Conceptualization, YL and HH; literature search, YL; data preparation & analysis, YL, HQQ, CX, FDM, KN, and HQ; data interpretation, YL, HQQ, CX, KN, XW, JG, AHS, DW, and HH; original draft writing, YL; review and revision, YL, HQQ and HH.

ACKNOWLEDGEMENTS

The sequencing data were provided through the Gabriella Miller Kids First Pediatric Research Program consortium (Kids First), supported by the Common Fund of the Office of the Director of the National Institutes of Health (www.commonfund.nih.gov/KidsFirst), awarded to Center for Applied Genomics.

CONFLICT OF INTERESTS STATEMENT

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

FUNDING INFORMATION

The study was supported by Institutional Development Funds from the Children's Hospital of Philadelphia to the

Center for Applied Genomics, The Children's Hospital of Philadelphia Endowed Chair in Genomic Research to HH.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

We confirm that all methods were carried out in accordance with relevant guidelines and regulations and all experimental protocols were approved by the Children's Hospital of Philadelphia (CHOP) Institutional Review Board (IRB 16-013278). Informed consent was obtained from all subjects or, if subjects were under 18, from a parent and/or legal guardian with assent from the child if 7 years or older. All necessary patient/participant consent has been obtained, and the appropriate institutional forms have been archived.

CONSENT FOR PUBLICATION

Not applicable.

DATA AVAILABILITY STATEMENT

The KidFirst data could be accessed at Kids First Data Resource Portal (DRC) (https://portal.kidsfirstdrc.org/login).

Yichuan Liu[1]
Hui-Qi Qu[1]
Xiao Chang[1]
Frank D Mentch[1]
Haijun Qiu[1]
Kenny Nguyen[1]
Xiang Wang[1]
Amir Hossein Saeidian[1]
Deborah Watson[1,2]
Joseph Glessner[1,2]
Hakon Hakonarson[1,2,3,4,5]

[1]*Center for Applied Genomics (CAG), Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA*
[2]*Department of Pediatrics, The Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA*
[3]*Division of Human Genetics, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA*
[4]*Division of Pulmonary Medicine, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA*
[5]*Faculty of Medicine, University of Iceland, Reykjavik, Iceland*

Correspondence
Hakon Hakonarson M.D., Ph.D. and Yichuan Liu Ph.D., Center for Applied Genomics, Children's Hospital of Philadelphia, 3615 Civic Center Blvd, Abramson Building, Philadelphia, PA 19104, USA.
Email: hakonarson@email.chop.edu and liuy5@email.chop.edu

ORCID

Yichuan Liu https://orcid.org/0000-0003-2023-072X
Hui-Qi Qu https://orcid.org/0000-0001-9317-4488

REFERENCES

1. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. CA Cancer J Clin. 2022;72(1):7–33.
2. Aprile G, Pisa FE, Follador A, Foltran L, De Pauli F, Mazzer M, et al. Unplanned presentations of cancer outpatients: a retrospective cohort study. Support Care Cancer. 2013;21(2):397–404.
3. May P, Garrido MM, Aldridge MD, Cassel JB, Kelley AS, Meier DE, et al. Prospective Cohort Study of Hospitalized Adults With Advanced Cancer: Associations Between Complications, Comorbidity, and Utilization. J Hosp Med. 2017;12(6):407–13.
4. Ahmed H, Carmody JB. On the Looming Physician Shortage and Strategic Expansion of Graduate Medical Education. Cureus. 2020;12(7):e9216.
5. Ali H, Harting R, de Vries R, Ali M, Wurdinger T, Best MG. Blood-Based Biomarkers for Glioma in the Context of Gliomagenesis: A Systematic Review. Front Oncol. 2021;11:665235.
6. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010;38(16):e164.
7. Wang J, Ma Z, Carr SA, Mertins P, Zhang H, Zhang Z, et al. Proteome Profiling Outperforms Transcriptome Profiling for Coexpression Based Gene Function Prediction. Mol Cell Proteomics. 2017;16(1):121-34.
8. Nusse R. Wnt signaling and stem cell control. Cell Res. 2008;18(5):523-7.
9. Latour M, Her NG, Kesari S, Nurmemmedov E. WNT Signaling as a Therapeutic Target for Glioblastoma. Int J Mol Sci. 2021;22(16):8428.
10. Kim SH, Lim KH, Yang S, Joo JY. Long non-coding RNAs in brain tumors: roles and potential as therapeutic targets. J Hematol Oncol. 2021;14(1):77.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.