Raising the Bar for Traumatic Brain Injury Biomarker Research: Methods Make a Difference

Linda Papa^{1,2} and Kevin K.W. Wang^{3,4}

Dear Editor:

Research in blood-based traumatic brain injury (TBI) biomarkers has exploded over the last two decades and continues to grow at an unwavering pace. Commercialization of these biomarkers is ongoing, with a number of companies seeking United States Food and Drug Administration (FDA) approval to market their tests for clinical applications. Although there are an abundance of articles being published, many lack the rigorous methods and reporting required to adequately evaluate these markers for clinical use. Too often, there are inadequate sample sizes, inappropriate control groups and outcome measures, variable definitions of TBI, and differences in when samples are drawn relative to time of injury. Importantly, variability in how samples are processed and performance characteristics of the assays themselves are rarely known or incompletely described. It is time to raise the bar for how blood-based TBI biomarker research is conducted, and to encourage researchers, laboratories, and companies (producing and/or running assays) to perform at a higher standard.

In the recently published study by Posti and coworkers, the performance of glial and neuronal biomarkers glial fibrillary acidic protein (GFAP) and ubiquitin C-terminal hydrolase-L1 (UCH-L1) was compared between CT negative mild TBI patients (based on American Congress of Rehabilitation Medicine [ACRM] criteria)¹ and orthopedic control patients.² Proteomic analyses of GFAP and UCH-L1 were conducted at Randox Laboratories Ltd. (Crumlin, County Antrim, United Kingdom) with Randox Biochip technology. Blood samples were drawn, when available, on days 1, 2, 3, and 7 after patients' admission to the hospital, and at a late time point somewhere between 3 and 10 months after the injury. The authors found no significant differences in the concentrations of GFAP or UCH-L1 between mild TBI patients with a negative CT ($n=55$) and orthopedic control patients ($n=44$).

This is in contrast to the findings from previous studies comparing orthopedic controls to mild TBI patients (also based on ACRM definition)¹ with and without intracranial lesions on CT.^{3–6} In a recently published study in *JAMA Neurology*, which enrolled 584 trauma patients within 4 h of injury, GFAP ($p<0.001$) and UCH-L1 ($p=0.049$) were both significantly higher in CT nega-

tive mild TBI patients than in patients with orthopedic injuries.³ Similar findings were published for GFAP in 2014.⁶ The question is why the study by Posti and coworkers did not show this difference using the same biomarkers. We will explore some possible explanations.

In the study by Posti and coworkers there was significant variability in the timing of samples over the first 24 h of admission, with >40% of enrolled patients lacking blood samples for analysis within the initial post-injury period. Only 55/93 (56%) of the mild TBI and 44/73 (60%) of the orthopedic controls had samples available within 24 h. Compounding this, sample time was based on admission to hospital and not on time of injury (which may have occurred long before presenting to the hospital). These factors can have a significant impact on the performance of a biomarker whose peak is very early after injury. Previous work in a large cohort of trauma patients found that UCH-L1 peaked within 8 h of injury and that GFAP peaked within 20 h.³ Perhaps the samples were not drawn within the optimal window of detection. This may be particularly important in milder injuries where smaller amounts of the biomarkers may be released after injury. Understanding the temporal profile of a biomarker is crucial to measuring it at the appropriate time.

With cases of mild TBI and orthopedic trauma, verification of symptoms and symptom burden must be carefully and prospectively performed by the study team/clinician (as done in other studies^{3–6}) and not simply obtained retrospectively from the patients' medical records. It is not uncommon for patients with orthopedic injuries to have a concomitant undiagnosed mild TBI, particularly with upper limb fractures.⁷ If symptoms are not carefully and prospectively assessed, patients will not be correctly classified. Curiously, the MRI reports of 30% of the orthopedic controls in the Posti study showed some kind of abnormality. The authors reported the abnormalities as not related to the TBI. However, one subject had evidence of a previous contusion, and five others demonstrated nonspecific white matter lesions. The question arises as to whether these lesions could have affected the biomarker levels. Next, the question of whether CT negative mild TBI patients had their symptoms adequately verified and how significant these suspected brain injuries actually were, arises. There is no description or quantification of loss of consciousness,

¹Department of Emergency Medicine, Orlando Regional Medical Center, Orlando, Florida.

²Department of Neurology and Neurosurgery, McGill University, Montreal, Quebec, Canada.

³Departments of Psychiatry and Neuroscience, and ⁴Program for Neurotrauma, Neuroproteomics and Biomarkers Research, University of Florida, Gainesville, Florida.

amnesia, confusion, or even a listing of Glasgow Coma Scale (GCS) scores among the mild TBI cases. A feature of the study by Posti and coworkers that could have potentially addressed this issue was not exploited. The extensive MRI imaging battery (T13D, T2, fluid-attenuated inversion recovery [FLAIR], susceptibility weighted imaging [SWI], diffusion-weighted imaging [DWI], and diffusion tensor imaging [DTI] sequences) was performed on 52 orthopedic control patients, but only on 4 patients with mild TBI. It is unfortunate that more patients with mild TBI did not undergo MRI imaging. This could have provided some unique insights into injury in the mild TBI group.

Distinguishing patients with mild TBI and concussion from other trauma, such as orthopedic injuries, is an excellent end-point. However, the most critical end-points in the acute care setting are detecting traumatic intracranial lesions on CT and identifying those who will require neurosurgical intervention. Recent and independent studies have shown the utility of GFAP and UCH-L1 in detecting traumatic CT lesions^{3-6,8-10} and in predicting neurosurgical intervention.³⁻⁵ In all of these studies, the biomarkers were run using chemiluminescent-based sandwich ELISAs (Banyan Biomarkers Inc.). A number of these studies also show progressive increases of these biomarkers relative to severity of injury from controls to those with a clinical diagnosis of TBI to those with CT abnormalities and those requiring neurosurgical intervention.³⁻⁵ Taken together, all of the published studies on the potential utility of TBI biomarkers in mild TBI, particularly GFAP and UCH-L1, remain very promising.

Conducting TBI biomarker research is no easy task, and requires considerable effort to enroll a sufficient number of patients to draw a valid conclusion. The study by Posti and coworkers, although within the range of many pilot studies in the field, only had an effective sample size of 99 patients (out of 166) in the first 24 h and decreased significantly thereafter over later time points. When a study is underpowered the validity of the results are always in question. Future TBI biomarker validation studies should make it a priority to justify sample size through *a priori* power calculations, otherwise they will suffer from “power failure.”¹¹

Among the most provocative explanations for the negative findings in the article by Posti and coworkers lies in the performance of the assay used to measure GFAP and UCH-L1. The precision, analytic sensitivity, and specificity of the assay being tested may have been reasons behind the discrepancy between this and other studies evaluating these two markers. Assay differences to quantify calcium-binding S100 β have been previously reported using two different analytical methods. Muller and coworkers¹² compared serum S100 β levels from the Liaison Sangtec 100TM (enzyme-linked immunosorbent assay [ELISA]) to the Elecsys S100TM (Roche Diagnostics, Mannheim, Germany) in patients with subarachnoid hemorrhage and TBI and found that S100 β concentrations were markedly different between the two assays. Accordingly, there are assay differences to be noted in Posti and coworkers’ study. The UCH-L1 and GFAP sandwich ELISA assays from Banyan Biomarker Inc. are run on a 96 well plate format (antigen protein standards are always co-run with the actual clinical samples), whereas the Randox Biochip technology runs the GFAP and UCH-L1 on a chip-based assay with the antibody-antigen sandwich complex detected and quantified using a CCD camera with a calibration curve used (off-line) to calculate the concentration of analytes. Although assay performance characteristics of these two biomarkers (such as inter-assay and intra-assay coefficients of variance, lower limit of detection, and lower limit of quantification) have been reported on the former assays,³⁻⁵ they have not all been reported for the latter platform.^{2,13}

Lower limits of quantification and detection (the ability to measure very low concentrations of a biomarker) are important when attempting to distinguish between very mild injuries. If the performance of the assays in the studies by Papa and coworkers³ and Posti and coworkers² are compared, there are differences in the lower limits of quantification (LLOQ) and detection (LLOD) between the Banyan and the Randox assays, respectively. In Papa and coworkers’ study,³ the LLOQ for the Banyan GFAP assay was 0.030 ng/mL (30 pg/mL) with an LLOD of 0.008 ng/mL (8 pg/mL). In contrast, the Posti and coworkers study² reported an LLOQ of 0.300 ng/mL (300 pg/mL) for the Randox GFAP assay, with no LLOD described. Correspondingly, the LLOQ for the Banyan UCH-L1 assay was 0.100 ng/mL (100 pg/mL) with an LLOD of 0.045 ng/mL (45 pg/mL)³ versus an LLOQ of 0.160 ng/mL (160 pg/mL) for the Randox UCH-L1 assay,² with no LLOD described. It is important to determine the significance of this. If one were to apply the Randox GFAP lower limit (0.300 ng/mL) to the population used in Papa and coworkers’ study,³ 87% of the CT negative mild TBI patients would have had readings below this level and, therefore, the results would have been almost indiscernible from those of the orthopedic control patients (98% of which were below this level). Further, this lower level (0.300 ng/mL) was higher than that in 40% of the CT positive patients with intracranial lesions included in Papa and coworkers’ study and, therefore, would not have been sensitive enough to detect many significant acute injuries.³ Therefore, limits of quantification and detection are important parameters that need to be determined during assay development. Assigning a value of zero to values that are below the level of quantification or detection is inaccurate and misleading.

Reasons for lack of concordance between different assays (for any biomarker) include: lack of standardization of the calibrating materials, differences in the specificity of the antibodies used, differences in the platforms used to measure the analytes, variability in forms of the biomarkers and the reactivity of the antibodies to these forms, and differences in the analytic performance of assays with particular reference to analytic sensitivity and assay imprecision and reference standard/standard curve utilization. Evidently, the current lack of industry standards and reporting requirements for commercial and research assays for TBI is a major obstacle to bedside translation. It is hoped, in time, that all TBI biomarker assays will be standardized and will exhibit comparable performance, as is the case for the majority of other clinical laboratory analytes today, especially where FDA approved assay platforms exist.¹⁴

In the search for the “perfect” diagnostic TBI biomarker/s, one must recognize that it is challenging to find a single protein or metabolite that is entirely “brain” specific. There is tremendous redundancy in the uses and functions of the, ~20,000 proteins that our genome encodes. Moreover, the body’s neural network is very sophisticated. Various cell types found in the brain, such as neurons and astrocytes, share structural and functional similarities with cells in other parts of the body, including those in the spinal cord, peripheral nerves, and solid organs.^{15,16} The degree to which they would contribute to release of GFAP and UCH-L1 after trauma will require further elucidation. Future studies will have to incorporate control groups with different organ injuries in order to evaluate release of potential biomarkers from different organs and gain understanding of the biomarker’s tissue distribution, subcellular locations, and release patterns into blood after trauma.

There is no biomarker for TBI that, to date, has met the rigorous standards of the FDA. If we do not adopt such rigorous standards in evaluating potential TBI biomarkers, the quest for a clinically

useful TBI biomarker will never be realized, and individuals with a brain injury will not benefit from the potential improvements in their care.

Controversies in research help educate and create opportunities to explore disparities among studies and inform the design, methods, and analysis of future clinical studies. If TBI biomarker research is to translate into clinical care, research methods need to be more robust and held to a higher standard. Otherwise, research studies remain academic exercises complicated by unnecessary controversies generated by flawed methodologies. Without strong evidence, clinicians will not accept or adopt such a blood test into their routine practice. Moving forward, more open and formal discussions must be initiated to shape and set guidelines for researchers and companies doing TBI biomarker work. Such efforts will more reliably and effectively convert research findings to the clinic, pre-hospital setting, playing field, battlefield, and bedside, where they can positively impact the lives of TBI patients and their families.

Acknowledgments

The authors give special thanks to (in alphabetical order) Gretchen M. Brophy, PharmD; Dallas C. Hack, MD, MPH; Lawrence M. Lewis, MD; Kara E. Schmid, PhD; Frank C. Tortella, PhD; and Robert D. Welch, MD, MS for critically reviewing this manuscript and for sharing their insight and expertise.

Author Disclosure Statement

Dr. Papa is an unpaid scientific consultant of Banyan Biomarkers, Inc. but receives no stocks or royalties from the company and will not benefit financially from this publication. Dr. Wang owns stock and receives royalties from Banyan Biomarkers Inc.

References

- Kay, T., Harrington, D.E., Adams, R., Anderson, T., Berrol, S., and Cicerone, K. (1993). American Congress of Rehabilitative Medicine; definition of mild traumatic brain injury. *J. Head Trauma Rehabil.* 8, 86–87.
- Posti, J.P., Hossain, I., Takala, R.S., Lienes, H., Newcombe, V., Outtrim, J., Katila, A.J., Frantzen, J., Ala-Seppala, H., Coles, J.P., Kyllonen, A., Maanpaa, H.R., Tallus, J., Hutchinson, P.J., van Gils, M., Menon, D.K., and Tenovuo, O. (2017). Glial fibrillary acidic protein and ubiquitin c-terminal hydrolase-L1 are not specific biomarkers for mild CT-negative traumatic brain injury. *J. Neurotrauma*. [Epub ahead of print.]
- Papa, L., Brophy, G.M., Welch, R.D., Lewis, L.M., Braga, C.F., Tan, C.N., Ameli, N.J., Lopez, M.A., Haeussler, C.A., Mendez Giordano, D.I., Silvestri, S., Giordano, P., Weber, K.D., Hill-Pryor, C., and Hack, D.C. (2016). Time course and diagnostic accuracy of glial and neuronal blood biomarkers GFAP and UCH-L1 in a large cohort of trauma patients with and without mild traumatic brain injury. *JAMA Neurol.* 73, 551–560.
- Papa, L., Lewis, L.M., Falk, J.L., Zhang, Z., Silvestri, S., Giordano, P., Brophy, G.M., Demery, J.A., Dixit, N.K., Ferguson, I., Liu, M.C., Mo, J., Akinyi, L., Schmid, K., Mondello, S., Robertson, C.S., Tortella, F.C., Hayes, R.L., and Wang, K.K. (2012). Elevated levels of serum glial fibrillary acidic protein breakdown products in mild and moderate traumatic brain injury are associated with intracranial lesions and neurosurgical intervention. *Ann. Emerg. Med.* 59, 471–483.
- Papa, L., Lewis, L.M., Silvestri, S., Falk, J.L., Giordano, P., Brophy, G.M., Demery, J.A., Liu, M.C., Mo, J., Akinyi, L., Mondello, S., Schmid, K., Robertson, C.S., Tortella, F.C., Hayes, R.L., and Wang, K.K. (2012). Serum levels of ubiquitin C-terminal hydrolase distinguish mild traumatic brain injury from trauma controls and are elevated in mild and moderate traumatic brain injury patients with intracranial lesions and neurosurgical intervention. *J. Trauma Acute Care Surg.* 72, 1335–1344.
- Papa, L., Silvestri, S., Brophy, G.M., Giordano, P., Falk, J.L., Braga, C.F., Tan, C.N., Ameli, N.J., Demery, J.A., Dixit, N.K., Mendes, M.E., Hayes, R.L., Wang, K.K., and Robertson, C.S. (2014). GFAP outperforms S100beta in detecting traumatic intracranial lesions on computed tomography in trauma patients with mild traumatic brain injury and those with extracranial lesions. *J. Neurotrauma* 31, 1815–1822.
- Jodoin, M., Rouleau, D.M., Charlebois-Plante, C., Benoit, B., Leduc, S., Laflamme, G.Y., Gosselin, N., Larson-Dupuis, C., and De Beaumont, L. (2016). Incidence rate of mild traumatic brain injury among patients who have suffered from an isolated limb fracture: upper limb fracture patients are more at risk. *Injury* 47, 1835–1840.
- Diaz-Arrastia, R., Wang, K.K., Papa, L., Sorani, M.D., Yue, J.K., Puccio, A.M., McMahon, P.J., Inoue, T., Yuh, E.L., Lingsma, H.F., Maas, A.I., Valadka, A.B., Okonko, D.O., Manley, G.T., Casey, I.S., Cheong, M., Cooper, S.R., Dams-O'Connor, K., Gordon, W.A., Hricik, A.J., Menon, D.K., Mukherjee, P., Schnyer, D.M., Sinha, T.K., and Vassar, M.J. (2014). Acute biomarkers of traumatic brain injury: relationship between plasma levels of ubiquitin C-terminal hydrolase-L1 and glial fibrillary acidic protein. *J. Neurotrauma* 31, 19–25.
- Welch, R.D., Ayaz, S.I., Lewis, L.M., Unden, J., Chen, J.Y., Mika, V.H., Saville, B., Tyndall, J.A., Nash, M., Buki, A., Barzo, P., Hack, D., Tortella, F.C., Schmid, K., Hayes, R.L., Vossough, A., Sweriduk, S.T., and Bazarian, J.J. (2016). Ability of serum glial fibrillary acidic protein, ubiquitin C-terminal hydrolase-L1, and S100B to differentiate normal and abnormal head computed tomography findings in patients with suspected mild or moderate traumatic brain injury. *J. Neurotrauma* 33, 203–214.
- Lewis, L.M., Schloemann, D., Papa, L., Fucetola, R., Bazarian, J., Lindburg, M., and Welch, R. (2017). Utility of serum biomarkers in the diagnosis and stratification of mild traumatic brain injury. *Acad. Emerg. Med.* [Epub ahead of print].
- Button, K.S., Ioannidis, J.P., Mokrysz, C., Nosek, B.A., Flint, J., Robinson, E.S., and Munafò, M.R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nat. Rev. Neurosci.* 14, 365–376.
- Muller, K., Elverland, A., Romner, B., Waterloo, K., Langbakk, B., Unden, J., and Ingebrigtsen, T. (2006). Analysis of protein S-100B in serum: a methodological study. *Clin. Chem. Lab. Med.* 44, 1111–1114.
- Posti, J.P., Takala, R.S., Runtti, H., Newcombe, V.F., Outtrim, J., Katila, A.J., Frantzen, J., Ala-Seppala, H., Coles, J.P., Hossain, M.I., Kyllonen, A., Maanpaa, H.R., Tallus, J., Hutchinson, P.J., van Gils, M., Menon, D.K., and Tenovuo, O. (2016). The levels of glial fibrillary acidic protein and ubiquitin C-terminal hydrolase-L1 during the first week after a traumatic brain injury: correlations with clinical and imaging findings. *Neurosurgery* 79, 456–464.
- Love, S.A., Sandoval, Y., Smith, S.W., Nicholson, J., Cao, J., Ler, R., Schulz, K., and Apple, F.S. (2016). Incidence of undetectable, measurable, and increased cardiac troponin I concentrations above the 99th percentile using a high-sensitivity vs a contemporary assay in patients presenting to the emergency department. *Clin. Chem.* 62, 1115–1119.
- Sofroniew, M.V., and Vinters, H.V. (2010). Astrocytes: biology and pathology. *Acta Neuropathol.* 119, 7–35.
- Bishop, P., Rocca, D., and Henley, J.M. (2016). Ubiquitin C-terminal hydrolase L1 (UCH-L1): structure, distribution and roles in brain function and dysfunction. *Biochem. J.* 473, 2453–2462.

Address correspondence to:

Linda Papa, MD, CM, MSc
Department of Emergency Medicine
Orlando Regional Medical Center
86 W. Underwood (S-200)
Orlando, FL 32806

E-mail: lpstat@aol.com