EDITORIAL

Noninvasive biomarkers: are we there yet?

Jeffrey Ojemann, MD

Department of Pediatric Neurosurgery, Seattle Children's Hospital, Seattle, Washington

As part of an ongoing focused effort to identify noninvasive biomarkers for pediatric brain tumors, Pricola Fehnel et al. have identified a combination of markers that appear specific to juvenile pilocytic astrocytomas (JPAs).⁷ By measuring levels of tissue inhibitor of metalloproteinase 3 (TIMP3) and basic fibroblast growth factor (bFGF) normalized to total protein, a profile was identified that distinguished patients with JPA from controls. Specifically, elevated TIMP3 beyond the control range characterized JPA (and other pathologies, especially medulloblastoma), whereas elevated bFGF was specific to JPA. Twenty-one patients with JPA contributed to this cohort and, in a smaller subset of 9 patients, follow-up levels of both TIMP3 and bFGF were noted to drop with treatment. Finally, another biomarker, matrix metalloproteinase 13 (MMP13), was elevated in medulloblastoma and not JPA.

This work represents an exciting direction in the development of tumor biomarkers. The authors imagine a day where long travel, sedated MRI, and the corresponding expense are replaced by an outpatient assay. These tests are practical. Each assay can be run on as little as 100 μ l of specimen. As with any screening test, the consequences of sensitivity and specificity shortcomings can be amplified and can limit the desirability of applying such a test to a larger population.

These markers certainly are related to tumor burden, although the exact link is unclear. Cultures from 3 tumors demonstrated TIMP3 (in the cytoplasm) and bFGF (in the nuclei), giving validation to the biomarker. MRI volume did not correlate with these levels so the exact link between tumor burden and levels remains unknown. Tumor treatment correlated with a decrease in urinary marker levels at the group level, but we are not given the individual data, and there were no instances of tumor recurrence. This is the key piece missing in achieving the goal of a biomarker. Whether all types of JPA will have similar profiles and show similar natural history in the evolution of urinary levels also remains unstudied. In fact, it will be

impossible to study these questions in single-institution series as the sample sizes are quite small. This total of 9 cases of pre- and postoperative levels was accumulated over 7 years, indicating how long it could take to have an adequately powered series to answer critical questions.

The authors note that a noninvasive marker could be especially valuable in the case of a chronic, benign condition such as JPA. What is an acceptable false negative in this setting? Even the role of delayed, sustained imaging of JPA beyond a few years is debated in the setting of gross-total resection, making the ideal sensitivity of a noninvasive test all the more unclear.^{1,5} A small residual may even regress, rather than grow, making the necessity of frequent surveillance appear less urgent.^{2,6} The consequence of a false positive could be additional imaging that might otherwise be considered unindicated in an asymptomatic patient. Even a low false-positive rate could eliminate the economic advantage of this test if a limited postresection surveillance imaging approach is the alternative. Conversely, compared with MRI, a false negative would delay identification of a recurrence, possibly postponing the identification of an easily treatable recurrence until the point that symptoms progressed and treatment has higher morbidity. Given long-term estimates of tumor recurrence of as low as 5%,⁴ it will be very difficult to accumulate enough patients to show an economic advantage to a urinary biomarker surveillance approach suggested by the authors.

At the risk of lessening the authors' enthusiasm, it may be worth looking at the history of other screening tests, such as that for prostate-specific antigen (PSA). In this test, a serum marker appeared to correlate well with tumor detection and following therapy. However, widespread use of screening PSA, across different populations, led to false positives and overtreatment with associated harm that led to screening recommendations much more limited than originally proposed.³ In parallel, it is reasonable to be concerned that the widespread dissemination of urine-based testing for tumor markers could lead to overtreatment. The

receiver operating characteristic curves reported for these markers are suggestive of a good test but the sample sizes are too small to be confident that a test that would likely be used routinely and frequently would not lead to a fountain of unnecessary imaging.

One great strength of the approach is that the test is on a waste product (urine). There is guarantee of a specimen in all cases and it can be collected by any patient, no matter where the patient received care. A national effort to accumulate data would be a herculean logistical task but, once set up, data would flow unencumbered. As with so many of our clinical questions, this multiinstitutional approach appears to be mandatory to determine whether a screening test can ever be developed for relatively low-incidence tumors, such as characterize the problems we face in our field.

<http://thejns.org/doi/abs/10.3171/2016.2.PEDS15724>

References

1. Dodgshun AJ, Maixner WJ, Hansford JR, Sullivan MJ: Low rates of recurrence and slow progression of pediatric pilocytic astrocytoma after gross-total resection: justification for reducing surveillance imaging. **J Neurosurg Pediatr** 17:569–572, 2016
2. Due-Tønnessen BJ, Helseth E, Scheie D, Skullerud K, Aamodt G, Lundar T: Long-term outcome after resection of benign cerebellar astrocytomas in children and young adults (0–19 years): report of 110 consecutive cases. **Pediatr Neurosurg** 37:71–80, 2002
3. Ilic D, Neuberger MM, Djulbegovic M, Dahm P: Screening for prostate cancer. **Cochrane Database Syst Rev** 31:CD004720, 2013
4. Johnston DL, Keene D, Bartels U, Carret AS, Crooks B, Eisenstat DD, et al: Low grade astrocytoma in children under the age of three years: a report from the Canadian pediatric brain tumour consortium. **J Neurooncol** 124:95–100, 2015
5. Ogiwara H, Bowman RM, Tomita T: Long-term follow-up of pediatric benign cerebellar astrocytomas. **Neurosurgery** 70:40–48, 2012
6. Palma L, Celli P, Mariottini A: Long-term follow-up of childhood cerebellar astrocytomas after incomplete resection with particular reference to arrested growth or spontaneous tumour regression. **Acta Neurochir** 146:581–588, 2004
7. Pricola Fehnel K, Duggins-Warf M, Zurakowski D, McKee-Proctor M, Majumder R, Raber M, et al: Using urinary bFGF and TIMP3 levels to predict the presence of juvenile pilocytic astrocytoma and establish a distinct biomarker signature. **J Neurosurg Pediatr** [epub ahead of print June 17, 2016. DOI: 10.3171/2015.12.PEDS15448]

Disclosures

The author reports no conflict of interest.

Response

Edward R. Smith, MD

Department of Neurosurgery, Boston Children's Hospital, Boston, Massachusetts

We read with interest the commentary on our article by Dr. Ojemann and appreciate his review. In short, we agree with nearly everything he says. In addition to sharing our enthusiasm for a potentially novel method of tumor detection, he highlights several valid limitations of this sort of

study: small patient population, potential concerns with diagnostic accuracy, and the need to proceed with caution when adopting new markers of disease, as illustrated by the history of PSA. Each of the points is valid, and merits response.

Regarding the small number of patients, we unequivocally agree that larger trials are needed to validate the utility of this, or any, marker. The purpose of this initial analysis was a proof-of-principle study. While it took time to obtain the specimens in this work from our center, the unique advantage of using urine means that future collection should be much easier. We have a precedent with an ongoing multicenter study for urinary biomarkers for pontine glioma in which samples are mailed from 12 hospitals nationally. With this approach, we have rapidly accrued both initial and longitudinal patient samples and we hope to now capitalize on this experience to expand our work to other tumor types, including JPA, and medulloblastoma.¹

This leads directly to the second point concerning diagnostic accuracy and how the level of accuracy might influence clinical practice. The clinical presentation and typical natural history of JPA presents an extreme end of the spectrum for brain tumor monitoring (benign and slow growing). We contrasted JPA with medulloblastoma because there may be greater utility in other brain tumor types that would benefit from more frequent monitoring. To be clear, we are not currently advocating this technique as a screening tool for the general population; rather, we anticipate that the initial application would be best suited to follow-up evaluation of known tumors. Moreover, the use of urinary biomarkers is not meant to supplant MRI or existing protocols; our hope is that they potentially enhance them by adding a cheap, fast, easy method that relies on data biologically distinct from imaging.^{2,3}

The cautionary tale of PSA is a valid one and bears careful examination. Two important lessons were learned from the long study of PSA. First, as mentioned by Dr. Ojemann, its overuse as a general screening test exposed the lack of accuracy in diagnosis and led to overtreatment-related harm. As mentioned above, the future use of biomarkers—including urinary biomarkers for neurosurgical disease—needs to be prospectively tailored to specific uses (such as follow-up of a known tumor) and restricted from general screening use until far more substantive validation has been performed.⁴ Second, a powerful tool present in urinary biomarkers and absent in PSA is the capacity for multiplexing, i.e., combining several biomarker species to markedly increase specificity and sensitivity, as presented in our paper. Combining bFGF, TIMP3, and MMP13 as markers (already possible with current technology) substantially improves diagnostic accuracy. While there are many examples of “single-species” biomarkers that are clinically useful (prolactin, β -human chorionic gonadotropin, etc.), our expectation is that future biomarker development will rely heavily on combined markers as “fingerprints” to provide increasingly accurate disease signatures to help with clinical decision-making.

Our hope is that this work will spur study of and collaboration on novel methods to advance the diagnosis,

prognosis, and treatment of neurosurgical disease. We are grateful to the *Journal of Neurosurgery: Pediatrics* for publishing our work, to the editors (including Dr. Ojemann) for their thoughtful reviews, and to our colleagues, with whom we are excited to continue working toward our shared goal of improving the outcomes of children entrusted to our care.

References

1. Akino T, Han X, Nakayama H, McNeish B, Zurakowski D, Mammoto A, et al: Netrin-1 promotes medulloblastoma cell invasiveness and angiogenesis, and demonstrates elevated expression in tumor tissue and urine of patients with pediatric medulloblastoma. **Cancer Res** **74**:3716–3726, 2014

2. Roy R, Yang J, Moses MA: Matrix metalloproteinases as novel biomarkers and potential therapeutic targets in human cancer. **J Clin Oncol** **27**:5287–5297, 2009
3. Smith ER, Zurakowski D, Saad A, Scott RM, Moses MA: Urinary biomarkers predict brain tumor presence and response to therapy. **Clin Cancer Res** **14**:2378–2386, 2008
4. Strimbu K, Tavel JA: What are biomarkers? **Curr Opin HIV AIDS** **5**:463–466, 2010