The need for validation of MI GPSai in patients with CUP: Comment on: "Machine learning analysis using 77,044 genomic and transcriptomic profiles to accurately predict tumor type" by J Abraham et al.

Cancers of unknown primary site (CUP) are a syndrome of many metastatic cancers without clinically detectable anatomical primary sites and represent 3-5% of all advanced cancers [1]. Although genetic and epigenetic alterations may partially explain the CUP syndrome and confound accurate molecular diagnosis, the majority of patients are eventually found to have very small invasive primary cancers from a variety of anatomical sites as documented by autopsies [2]. The nature of the metastasis in CUP and response to site specific treatment, once the tissue of origin is determined, appear very similar to those of patients with corresponding metastatic cancers from obvious or known primary sites [2,3]. The appropriate and precise treatment of patients with CUP depends on the determination of the tissue of origin or cancer type for each individual patient. This is now possible in most patients using immunohistochemical staining panels as well as molecular cancer classifiers. Next generation sequencing and comprehensive molecular profiling results should most often be interpreted after knowledge of the cellular context to guide more precise treatment selection. To this end, the MI GPSai (Caris Life Sciences, Phoenix, AZ) validation for known cancers was recently published [4]. However, several aspects of this validation may not apply to patients with CUP and leads to a larger discussion of how to adequately validate the accuracy of molecular testing to determine the tissue of origin.

A key fundamental aspect of targeted validation of diagnostic accuracy for the MI GPSai is that it needs to be performed in patients with CUP. In Abraham et al., validation of MI GPSai was conducted using specimens with known cancer types as opposed to those diagnosed with CUP, with 13,661 of the validation samples coming from known primary samples and only 1,107 CUP samples. Of the 12,699 known primary validation samples above the performance threshold, 56% (n=7087) were from primary tumors and 44% (n=5426) were from metastatic sites. Assay performance was decreased in poorly differentiated or metastatic cases, and diagnostic accuracy was ultimately determined by the ordering physician without consensus-based adjudication; thus, no independent verification was performed. While demonstrating accuracy in known cancers represents an important initial step, validation should be performed in patients with CUP before oncologists and pathologists can feel confident that the molecular diagnosis of the tissue of origin is accurate. Of the 1292 CUP cases reported, MI GPSai assigned a diagnosis in only 71.7% of cases and the accuracy of these diagnoses are questionable. Only 46 of 1292 cases (3.5%) had a score greater than 0.999. In 19 of these 46 cases the diagnoses changed following analysis with MI GPSai, but details were provided in only 2 anecdotal cases. Therefore, the clinical utility and accuracy of MI GPSai in patients with CUP requires additional validation.

Validation of the accuracy of any molecular cancer classifier can be difficult in patients with CUP short of comparing the molecular diagnosis with primary sites found at autopsy. There are 3 practical methods other than autopsies which can provide very strong circumstantial evidence of accuracy, as outlined previously in 171 CUP patients [2,3,5]. First, evaluation of those patients with CUP who subsequently develop clinically-detected primary sites many months after their initial presentation (latent primary sites) offers a direct method or gold standard other than autopsy to assess the accuracy of a molecular diagnosis. Using the example of another molecular cancer classifier, the 92-gene RT-PCR-based CancerTYPE ID (Biotheranostics Inc., San Diego, CA), latent primaries were documented in 24 patients and the initial molecular diagnosis assessed was compared to the subsequent identified primary site, with CancerTYPE ID demonstrating 75% accuracy [2,3,5,6].

Second, single diagnoses made simultaneously by IHC staining can be compared to molecular diagnoses. Forty of 52 patients (77%) with specific IHC diagnoses matched the CancerTYPE ID diagnoses. Molecular classifiers should also be blindly compared to IHC (the accepted historical diagnostic standard) in known cancers and others with challenging metastatic cases with poorly differentiated or undifferentiated tumors [7,8].

Third, additional directed IHC staining and clinical and histologic findings obtained after the molecular diagnoses can confirm the diagnoses. In 74% of cases (26 of 35), additional analysis supported the accuracy of CancerTYPE ID. The MI GPSai described by Abraham et al. has not had adequate validation in the very group of patients with CUP for which it was designed to diagnose the tissue of origin. It is possible that machine learning combined with next generation sequencing may be an effective approach for tumor classification; however, without stronger validation studies in patients diagnosed with CUP the performance and accuracy remain questionable.

CRediT authorship contribution statement

F. Anthony Greco: Conceptualization, Investigation, Writing – original draft, Writing – review & editing, Visualization.

Declaration of interests

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Dr. F. Anthony Greco is a medical advisor and member of the speaker's bureau for Biotheranostics, Inc.

References

[1] G.R. Varadhachary, J.L. Abbruzzese, R. Lenzi, Diagnostic Strategies for Unknown Primary Cancer, Cancer 100 (9) (2004) 1776–1785, doi:10.1002/cncr.20202.

[2] J.D. Hainsworth, F.A. Greco, Cancer of Unknown Primary Site: New Treatment Paradigms in the Era of Precision Medicine, in: American Society of Clinical Oncology Educational Book, 38, American Society of Clinical Oncology (ASCO), 2018, pp. 20–25.

[3] F.A. Greco, J.D. Hainsworth, Cancer of unknown primary site, in: DeVita Jr, T. Vincent, Theodore S. Lawrence, S.A. Rosenberg (Eds.), Principles and Practice of Oncology, 10th ed., Wolters Kluwer, Philadelphia, 2015, pp. 1720–1737.

[4] J. Abraham, A.B. Heimberger, J. Marshall, et al., Machine learning analysis using 77,044 genomic and transcriptomic profiles to accurately predict tumor type, Transl Oncol 14 (2021), doi:10.1016/j.tranon.2021.101016.

[5] F.A. Greco, W.J. Lennington, D.R. Spigel, J.D. Hainsworth, Molecular profiling diagnosis in unknown primary cancer: Accuracy and ability to complement standard pathology, J Natl Cancer Inst 105 (11) (2013) 782–790, doi:10.1093/jnci/djt099.

[6] F.A. Greco, D.R. Spigel, D.A. Yardley, M.G. Erlander, X.-J.X. Ma, J.D. Hainsworth, Molecular Profiling in Unknown Primary Cancer: Accuracy of Tissue of Origin Prediction, Oncologist 15 (5) (2010) 500–506, doi:10.1634/theoncologist.2009-0328.

[7] F.A. Greco, W.J. Lennington, D.R. Spigel, J.D. Hainsworth, Poorly Differentiated Neoplasms of Unknown Primary Site: Diagnostic Usefulness of a Molecular Cancer Classifier Assay, Mol Diagnosis Ther 19 (2) (2015) 91–97, doi:10.1007/s40291-015-0133-8.

[8] L.M. Weiss, P. Chu, B.E. Schroeder, et al., Blinded comparator study of immunohistochemical analysis versus a 92-gene cancer classifier in the diagnosis of the primary site in metastatic tumors, J Mol Diagnostics 15 (2) (2013) 263–269, doi:10.1016/j.jmoldx.2012.10.001.

F. Anthony Greco*

Sarah Cannon Research Institute, Nashville, TN, 32703, USA

Tennessee Oncology, Nashville, TN, 32703, USA

*Corresponding author. 250 25th Ave N, Nashville, TN, 32703, Phone +1 (615-289-3057), Fax +1 (615-320-1225).

E-mail address: fgreco@tnonc.com