Biomarker accuracy: exploring the truth

The American theologian, writer and minister, Tryon Edwards once said, “Accuracy of statement is one of the first elements of truth; inaccuracy is a near kin to falsehood.” The same statement could be substituted for analytical measurements. The real question is how to define ‘truth’. Is it recovery of a spiked sample using a known quantity of analyte? Is it agreement with another method? Or is it quantifying a set of samples within a manufacturer’s predefined range? Depending on your background you may prefer one approach over the others. What happens when two methods produce different results for the same target? How do you determine which assay is producing the correct answer? These are the questions that biomarker assay scientists face every day and why it is important. Biomarker accuracy is critical for defining the exposure–response relationship and in the clinical diagnostic environment. In the former, biomarker concentrations along with drug concentrations are used to construct an exposure–response relationship. The model is superimposed on the adverse events profile to define the therapeutic window. If the biomarker value is over- or under-estimated, efficacious responses would be incorrect and the subsequent exposure–response relationship could lead to faulty interpretation. In clinical diagnostics, most assays are defined by a cut-point, a concentration whereby patient treatment decisions are made. In that situation, accuracy is needed to ensure that patients with the disease are treated appropriately, and that normal individuals are not prescribed unneeded medication.

In the field of biomarkers, the ability to demonstrate accuracy is more challenging than in other analytical disciplines because the analytes are endogenous, heterogeneous and often structurally different from the calibrator. It is for these reasons that some individuals in the field question the validity of traditional bioanalytical accuracy testing for biomarkers. Another consideration for biomarkers is the association with diagnostics. In diagnostic medicine, one definition for accuracy is the ability of a test to correctly identify patients with a specific disease. As a result, manufacturers will provide samples with predefined values that can be tested to verify that the observed values agree with the acceptable limits. However, research use only or in-house developed assays lack proficiency samples, so different approaches must be used to assess accuracy. I also believe that accuracy is more than just trueness, but represents the true in vivo concentration at the time of collection. Given the complications associated with measuring biomarkers, I recommend additional experiments that can complement a traditional spiked recovery experiment for assessing accuracy.

Keywords: accuracy • biomarker • diagnostics • pre-analytical • spiked recovery

What am I measuring?
The first assumption of any method is that the results are accurate and the test measures the analyte of interest. The ICH guidance states, “The accuracy of an analytical procedure expresses the closeness of agreement between the value which is accepted either as a conventional true value or an accepted reference value and the value found. This is sometimes termed trueness.” Two fundamental assumptions in bioanalysis are that you understand what is being measured, and that the calibrator represents the target analyte. However, most protein biomarker assays use recombinant proteins from engineered systems, so the calibrator may not adequately represent the endogenous protein. Depending on the system, recombinant proteins may lack proper folding, post-translational modifications, or it may be a portion of the full-length molecule [1]. As a result, experiments that ensure the test system is capable of measuring the endogenous protein are essential. Over the years, I have applied ex vivo stimulation approaches to produce cytokines, phosphoproteins or growth factors at concentrations that are above the LOQ to mimic a disease condition [2]. For additional confidence, we routinely
demonstrated modulation of the target by inhibitors, agonist titration, or variable time of exposure [3]. The combination of these two experiments provides confidence that the method is capable of detecting the protein of interest in a biologically relevant system. Another important experiment that provides confidence in accuracy is orthogonal method comparison. If a gold-standard assay is available, then the samples can be split and tested in both systems and compared for agreement [4]. Another elegant experiment that provides definitive characterization of what is being measured is an affinity capture, protein separation and identification by MS [5]. These experiments are often difficult to obtain given the limited resources, organizational constructs, technical expertise and availability of the platforms. However, this experiment provides the tools to describe the molecular species that were captured by the immunoassay reagent. These experiments provide greater understanding of the test method and what the assay is measuring.

Did I introduce artifacts?
Another important consideration for overall accuracy is sample collection, processing and storage. Any of these steps can introduce an artifact into the final result. There are several choices of additives and anticoagulants that prevent proteolytic activity of the clotting factors, prevent platelet activation and provide a carbon source for whole blood samples. Depending on the anticoagulant selection, results can vary significantly [6]. Processing differences can also alter the final concentration. Platelets serve as reservoirs for many proteins found in circulation, so if the processing steps fail to account for the platelet contribution, then the results could be altered [7]. As a result, we recommend evaluating the effects of platelet-rich and platelet-free plasma to determine the best collection strategy. Storage is the final factor that can influence the result. One observation that I have made is that storage recommendations are assay specific for testing systems that use antibodies to measure a protein. The effects of storage can change the structure of the target protein and as a result change the antigen–antibody interaction leading to differences in immunoreactivity. Therefore, it is essential to characterize the storage stability using the test method and not apply literature values, or other methods for the same protein.

“...it is essential to characterize the storage stability using the test method and not apply literature values, or other methods for the same protein.”

In summary, I believe that biomarkers represent a competitive advantage when developed and applied appropriately. Accuracy is a fundamental assumption of bioanalysis. Unbiased results are needed in the clinical diagnostic field and in the area of PK/PD. Given the endogenous nature of these analytes and the lack of a true reference standard, additional experiments are needed to define accuracy. In addition, special consideration must be given to collection and storage to avoid pre-analytical artifacts. Finding laboratories with the correct expertise to conduct these experiments often requires working cross-functionally, or across organizations in the case of outsourced assays. These experiments require special expertise and resources; however, the return on investment can be significant. These efforts should start early in the drug-discovery paradigm. I believe that more careful examination of what is being measured will yield better results, better decisions and better claims in the label. I hope that as a community, we begin to recognize the importance of developing assays and reagents with a deeper understanding of what we are measuring.

Financial & competing interests disclosure
The author has no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.
References


