The National Academy of Clinical Biochemistry
Laboratory Medicine Practice Guidelines

Draft Guidelines

TUMOR MARKERS

Eleftherios P Diamandis, MD, PhD, FRCPC

Catharine Sturgeon, MRCPath, PhD, FACB

Barry Hoffman, PhD, FCACB

We present here for comment and constructive critical appraisal the first draft of the new NACB Guidelines on the use of Tumor Markers in the clinic. The preparation of these draft guidelines has been achieved with the expert input of the editors and members of the guidelines sub-committees. The material in these guidelines represents the opinions of the sub-committees and editors and does not represent the official position of the National Academy of Clinical Biochemistry. The National Academy of Clinical Biochemistry is the official academy of the American Association for Clinical Chemistry.

Development of the Draft NACB Guidelines

Principles and methods supporting the development of evidence-based guidelines in laboratory medicine have recently been comprehensively reviewed (1) and some of the particular difficulties related to development of guidelines for tumor marker use have been discussed (2). Where possible, within the constraints of practicability, a similar approach to those recommended in these papers has been taken in this first draft of the Guidelines presented here. We wish to emphasise that this second web posting represents only a further stage in their development, as over the next six months the views of relevant professional organizations and other interested parties will be actively sought and incorporated, prior to publication of the final version of this document.

These draft guidelines, which consider general quality requirements [Section 2] and the use of tumor markers in sixteen major malignancies [Sections 3A to 3P – fifteen of which are posted here, with one to follow shortly] have been developed by Sub-committees under the chairmanship of a recognized expert in the field. Sub-Committee chairmen and members are listed in the relevant section of this document.

As far as has been possible, the same general format has been used in each section, each of which includes tables summarizing the currently available tumor markers for the malignancy being considered, together with the phase of development for each marker as well as the level of evidence (LOE) for their clinical use. The levels of evidence grading system used is based on that described by Hayes *et al* [Table 1] (3). A second table in each section summarizes the NACB guidelines for the use of markers in each malignancy and compares these with recommendations from other representative guidelines.

Table 1. Levels of evidence (LOE) for grading clinical utility of tumor markers (3)

Level	Type of evidence
I	Evidence from a single high-powered prospective controlled study specifically designed to test the marker, or evidence from a meta-analysis, pooled analysis or overview of level II or III studies.
II	Evidence from a study in which marker data are determined in relationship to a prospective therapeutic trial that is performed to test a therapeutic hypothesis but not specifically designed to test marker utility.
III	Evidence from large prospective studies.
IV	Evidence from small retrospective studies.
V	Evidence from small pilot studies.

Clinical issues that enhance reliability of utility of tumor markers

[When reviewing these guidelines it is helpful to consider the following comments on some of the complex issues relating to evidence-based tumor marker use in the clinic. These have been kindly provided by a member of our Steering Committee, Professor Daniel Hayes, to whom we are most grateful.]

As with all diagnostic tests, tumor markers are surrogate indicators that can be used clinically to increase or decrease the clinician's suspicion that a future clinically important event, such as a new cancer, recurrence, progression, or death will or will not happen, and/or that a specific treatment will reduce that risk. Markers can be used to determine risk, screen for early cancers, establish diagnosis, estimate prognosis, predict that a specific therapy will work, or monitor for disease recurrence or progression (3). The value of tumor markers is that they permit more efficient application of therapies, which should result in applying the therapy to those patients most likely to benefit while reducing exposure to toxicities for those patients who would not benefit (4).

Markers are only useful if three circumstances pertain:

- The marker results are appropriate precisely for the situation at hand: risk, screening, diagnosis, prognosis, prediction, monitoring.

- The marker results separate patients into two or more populations whose outcomes differ so strikingly that they and their caregiver would treat one group differently than another. [This consideration depends on several factors, including the endpoint in question (patients might be more willing to accept therapy for very small mortality reductions but not for similar reductions in occurrence of a new cancer) the toxicity of the therapy (patients are more likely to accept a therapy with small benefits if the toxicities are few), and the cost of the therapy.]

- The estimate of the separation in outcomes for marker positive and negative is reliable.

These issues are inter-related. For example, studies of the prognostic value of a marker that do not consider the manner in which the study populations were treated are not helpful to the clinician trying to decide whether to apply treatment. Indeed, in breast cancer, one might conclude that HER2 over-expression is associated with a poor prognosis, a favorable prognosis, or not associated with either if one studied a patient population that had been variably treated in either the adjuvant or metastatic setting with different types of chemotherapies, different types of hormone therapies, and Trastuzumab (5). These variable conclusions might be reached since HER2 is a weak or moderately unfavorable prognostic factor in patients who receive no therapy, it appears to predict weakly or moderately for resistance to chemotherapy regimens that do not contain anthracyclines or taxanes, but it may predict for sensitivity to chemotherapy regimens that do have these, it appears to predict for resistance to selective estrogen receptor modulators like tamoxifen but for sensitivity to estrogen ablation strategies like aromatase inhibitors, and it is a very strong predictor of response and benefit from the anti-HER2 humanized monoclonal antibody, Trastuzumab.

Furthermore, while statistical analysis is, of course, important to estimate the reliability of how likely two marker-identified groups might be different, the p-value alone does not indicate clinical utility. If a study is sufficiently powered, a small difference in outcomes of two groups separated by marker results ("positive vs. negative") might be statistically significant. Too often an investigator will conclude that a marker is clinically useful because a derived p-value is <0.05. Rather, it is more important, for clinical utility, that one population (marker positive or negative) does extremely well while the other does very poorly, so that one group might accept the therapy of interest while the other would elect not to. In this case, it is imperative that the p-value does suggest statistical significance, but it is not the determining factor for clinical utility. Finally, a single study does not establish a scientific fact. Rather, secondary validation of the results of an interesting study in a subsequent data set is imperative, and the validation study should use the same assay, the same cut point(s), and, importantly, the same types of patients.

In summary, acceptance of a tumor marker for clinical utility requires careful and thoughtful study design so that the results are meaningful in the clinical setting. Unfortunately, most tumor marker investigations are studies of convenience, using archived samples that happen to be available (3). Such studies [LOE III] are useful to generate hypotheses, but as in all science, without careful investigational planning and design the results cannot be accepted as fact. Indeed, LOE II studies, in which the marker is considered prospectively as a secondary objective in a clinical trial, or better yet, LOE I studies in which the marker question is the primary objective, are much more likely to yield acceptable results. In other words, it is better to ask the question and get an answer, rather than to get an answer and then ponder the question. Such evidence-based considerations are particularly important when patient lives are at stake.

INVITATION FOR PEER REVIEW

We now invite all interested parties to comment on the contents of the individual sections of these guidelines, which are listed in the Table of Contents below. Information about important references inadvertently omitted *etc* will also be most welcome. As indicated within each Section, all comments should be addressed *via* e-mail to C.Sturgeon@ed.ac.uk, with copies to ediamandis@mtsinai.on.ca and the relevant Sub-Committee Chairman. We would like to thank you in advance for your highly valued and much appreciated input.

REFERENCES

1. Oosterhuis WP, Bruns DE, Watine J, Sandberg S, Horvath AR. Evidence-based guidelines in Laboratory Medicine: Principles and methods. Clin Chem 2004;50:806-18

2. Sturgeon C. Practice Guidelines for tumor marker use in the clinic. Clin Chem 2002;48:1151-59

3. Hayes DF, Bast R, Desch CE, Fritsche H, Kemeny NE, Jessup J, *et al.* A tumor marker utility grading system (TMUGS): A framework to evaluate clinical utility of tumor markers. J Natl Cancer Inst 1996;88:1456-66.

4. Hayes DF. Prognostic and predictive factors for breast cancer: translating technology to oncology. J Clin Oncol 2005;23:1596-7

5. Yamauchi H, Stearns V, Hayes DF. When is a tumor marker ready for prime time? A case study of c-erbB-2 as a predictive factor in breast cancer. J Clin Oncol 2001;19:2334-56

Table of Contents