Tumor Markers: Precision, Accuracy, Validity and Statistical Matters of Concern

Md Mahamudur Rahman[1]
Md Monoarul Haque[2*]
Md Mahmuduzzaman Mian[3]
Md Shah Jalalur Rahman Shahi[1]

[1]Department of Immunology
Bangladesh University of Health Sciences
Dhaka, Bangladesh.

[2]Division of Public Health
Bangladesh University of Health Sciences
Dhaka, Bangladesh.

[3]MNS Department
BRAC University
Dhaka, Bangladesh.

*Correspondence to:

Md Monoarul Haque
Fellow
Department of Community Nutrition
Faculty of Public Health
Bangladesh University of Health Sciences (BUHS)
125/1, Darus Salam, Mirpur, Dhaka-1216, Bangladesh
Cell: +8801915839550
E-mail: monoarmunna@yahoo.com

Abstract

Tumor markers are often very helpful to diagnose, staging, treatment planning and following up of the cancer patients. But still deluge of tumor markers often presents an unethical excess laboratory experience of the patients which is neither helpful nor necessary. So this is the appropriate time to think again about the accuracy of the tests to measure the parameters without error, the precision of the tests to measure the parameter they intend to measure. Our acceptance or rejection of a tumor marker testing must be guided by statistical validity of the test with an appropriate and knowledge based cut-off value determination. Statistical analysis of the tumor marker should be proportional to the empirical judgments of the clinicians. For the clinicians it is a duty to advise not only for all the tests needed for appropriate treatment but also advise only the tests those are direly needed. Otherwise tumor markers which should be an important weapon to fight against cancer will be another burden over the cancer patients. In this article we intended to focus on the technical and statistical side of different tumor marker test most concentrated around the breast cancer tumor markers. But still this review may be concerned relevant and universal for each and every cancer type and tumor markers.

Key words: Tumor marker; cut-off point; accuracy; precision.

INTRODUCTION

Healthcare providers are increasingly embracing the principles of evidence-based medicine, which essentially is "the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients"[1]. Appropriately formulated clinical guidelines or protocols can facilitate an evidence-based approach to medicine[2], and numerous guidelines have been developed across a wide range of clinical areas. Available guidelines address issues ranging from the broad (e.g., how to provide medical or surgical services) to the specific (e.g., which diagnostic or screening test to order). Accompanying a comprehensive discussion of the advantages and disadvantages of clinical guidelines, a useful overview of international activity in this area has recently been published[2,3].

Serum tumor marker (STM) estimation is often used in clinical practice in monitoring response to treatment and as a predictor of treatment failure and relapse. However, there are pitfalls in interpretation, particularly in the immediate post treatment period, when a rise in titre could be observed, the phenomenon being termed as "flare". A literature search was done to examine this phenomenon for some of the commonly used serum tumor markers in malignancies. This phenomenon has been documented with respect to AFP, beta HCG, CEA, AC 15.3, PSA, CA 19.9 and CA 125 with or without other evidence of progression. A practical approach would be to correlate serum tumor marker values with other clinical and radiological parameters, and not to rely exclusively on serum marker values to guide therapy.

Breast cancer and tumor markers

During the past few decades, with the explosion of molecular technology and understanding of the biology of breast cancer, numerous studies have been performed to identify prognostic and predictive factors in breast cancer, with mixed success. Multiple expert panels have convened to analyze available data in order to establish guidelines for the use of tumor markers, but their recommendations have been very conservative[4,5]. The American Society of Clinical Oncology (ASCO) convened a panel of experts that first published recommendations regarding the use of circulating and tissue-based tumor markers in breast cancer in 1996[6]. The ASCO panel evaluated multiple serum markers for breast cancer, including assays for MUC1 protein (cancer antigen [CA] 15-3 and CA 27.29), carcinoembryonic antigen (CEA), and the circulating extracellular domain of Her-2/neu. The panel did not recommend monitoring of any of these markers for screening, diagnosis, staging, or routine surveillance of patients free of detectable disease. Measurement of CA 15-3 or CA 27.29 and/or CEA was recommended, however, to monitor selected patients with Metastatic Breast Carcinoma (MBC) undergoing palliative therapy[7]. Routine measurement of multiple tissue markers was also discussed in the guidelines. The panel recommended routine measurement of estrogen and progesterone receptors (ER and PR, respectively) to identify patients most likely to benefit from endocrine therapy in either the early or metastatic disease settings. In addition, measurement of Her-2/neu over-expression and possibly amplification was recommended for all patients at the time of initial diagnosis or recurrence, as it is predictive of response to trastuzumab (Herceptin®; Genentech, South San Francisco, CA), a monoclonal antibody directed against the Her-2/neu receptor[8].

When a tumor marker is useful?

When evaluating tumor markers for use in clinical practice, clinicians should consider their utility, the magnitude of their effects, and their reliability. Tumor markers can be useful at multiple stages of cancer diagnosis and treatment (For example, for individuals who do not have cancer, a marker may be helpful in determining the risk for developing the disease and/or it may be beneficial for screening for disease. Once an abnormality is found, a tumor marker may be helpful for distinguishing between benign and malignant processes or between different malignant processes. After confirmation of a cancer diagnosis, tumor markers can help monitor disease status during and after therapy[9,10]. Tumor markers can also help determine prognosis independent of therapy and predict response to therapy. Prognostic factors reflect the metastatic potential and/or growth rate of the tumor and are used to select patient outcomes without consideration of treatment given[11]. Predictive factors, on the other hand, reflect the sensitivity or resistance of a tumor to a therapeutic agent and therefore are used to predict which patients are likely to respond to a specific treatment[12]. Once a tumor marker use has been identified, it is important to determine the magnitude of the difference in outcomes for that particular use between those who are marker positive and those who are not. By evaluating the difference in outcome, regardless of treatment, between a patient positive for a given prognostic factor and one who is negative for the factor, the relative strength of a prognostic factor can be determined[13]. For example, a breast cancer patient with disease in the lymph nodes at the time of diagnosis is two to three times more likely to have a breast cancer event (local recurrence or distant metastasis) than a patient without lymph node involvement, regardless of treatment. Since lymph node status has classically been used to make clinical decisions, we have arbitrarily designated it as a "strong" prognostic factor, using it as the gold standard to set the criteria for consideration of other, putative markers[14].

Predictive factors can also be classified as weak, moderate, or strong, depending on their ability to predict response to, and therefore benefit from, a given therapy. One measure to permit comparison of the relative strengths has been designated the "relative predictive value" (RPV), the ratio of the likelihood that a factor-positive patient will respond to treatment to the likelihood that a factor-negative patient will respond to treatment. As with prognostic factors, arbitrary classes of prediction factors were proposed for breast cancer therapies based on what has been accepted by consensus, in this case ER[15]. Adjuvant tamoxifen therapy has been shown to decrease recurrence rates for ER-positive patients by 40%–50%, whereas ER-negative patients obtain minimal, if any, benefit from hormonal therapy[16]. Therefore, the RPV is >8. Similarly, the majority of patients with ER-positive MBC have a clinical response to hormonal therapy, whereas patients with ER-negative disease do not respond. ER status is therefore a strong predictive factor for response to hormonal therapy[17, 18].

Precision and Accuracy of Tumor markers

Tumor marker is only useful if the estimate of its magnitude is reliable and reproducible. In this regard, many investigators conclude that their marker of interest has clinical utility if in their study the difference in outcomes between marker "positive" and marker "negative" patients is less than conventional measures of statistical significant (p < .05). This conclusion may be mistaken. Statistical significance only suggests that in the population chosen for that study, the differences observed are likely not to be a result of chance alone. It does not imply clinical utility, nor does a p-value <0.05 document the validity of the tumor marker. Although it is important to determine that the differences in outcome achieve statistical significance, statistical significance alone does not determine clinical utility.

In addition to determining when to use a tumor marker and the magnitude of its effect, it is important to ensure that the technical aspects of the marker are reliable and reproducible and that the study design and conduct are appropriate to test the marker for a clinical use of interest. Several problems with tumor marker studies, including technical, analytical, and trial design issues, have limited the introduction of new prognostic and predictive factors into routine clinical practice[9].

Analytical issues important to consider

Determination of assay results can also vary, even for a single type of assay. For example, with visual assays such as IHC for ER and Her-2/neu, intra- and inter observer variability leads to differences in interpretation[19]. Some attempts have been made to standardize interpretation, such as development of the so-called "Allred score" for semiquantitation of ER expression, but these have not been universally adopted. Automated and semiautomated systems appear to be highly accurate and are likely to be more reproducible[20].

Cut-off point determination

Regardless of the assay, one has to select some value or level that distinguishes positive from negative results. However, there is no consensus regarding correct methods to establish cutoff points, and different studies of the same prognostic or predictive factor can have widely varying "optimal" cutoff points[21]. Cutoff points may be defined using either arbitrary or data-derived methods. One approach is to consider any value greater than two standard deviations above the mean for normal subjects to be positive. Cutoff points can also represent arbitrary values within affected patients; for example, one might decide that 10%, 50%, or 90% of affected patients will be classified as "positive." Others have defined cutoff points based on technical factors, such as the limit of detection of the assay[17]. Deriving cutoff points based on patient outcome data may provide more accurate values. For example, the cutoff point for ER expression was first defined by limits of the assay and later by determining the optimal level that distinguished those patients who respond to hormonal therapy from those who do not. In another example, the cutoff point for the CellSearch™ assay for circulating tumor cells was initially determined based on differences in time to progression of a test set of patients with metastatic disease, and this cutoff was then validated with an independent but similar patient cohort from the same study[22].

Recently, a novel data-derived method to select cutoff points, designated subpopulation treatment effect pattern plot (STEPP) analysis has been proposed. STEPP analysis evaluates outcomes to specific treatments in sub-populations of patients within randomized clinical trials or meta-analyses[23].

Statistical Analysis

Of course, statistical analysis is necessary to determine that the observations are not a result of chance alone. However, once a tumor marker has been identified and validated, it is important to determine the relative value of the marker in the context of previously identified prognostic and/or predictive factors, such as lymph node status and tumor size, using some type of a multivariate analysis. If such an evaluation is not performed, clinicians will be unable to determine the usefulness of incorporating the new marker into routine clinical practice.

Real world clinical example

Her-2/neu

The first report of Her-2/neu as a prognostic factor in breast cancer was published in 1987[24]. Since then, more than 200 papers addressing this topic have been published, with widely mixed and disparate results[25]. Different authors have concluded that Her-2/neu is associated with poor outcomes, no difference in outcome, or even favorable outcomes. Indeed, a great deal of this confusion could have been avoided if the investigators would have addressed the components described above: (a) What is the intended use, (b) What is the magnitude of difference between positive and negative for that use, and (c) How reliable is the estimate of the magnitude?

Overall, studies support that Her-2/neu overexpression is a poor prognostic factor, although its magnitude appears weak. Its role as a prognostic factor thus remains unclear.

CONCLUSION

Tumor markers, when well defined, can play a significant role in prediction and prognosis for breast cancer patients. Because of the abundance of poorly designed tumor marker studies to date, however, very few markers have been accepted for routine use by groups such as ASCO. When designing studies to establish a new tumor marker, or new use for an old marker, it is important to address the utility, magnitude, and reliability of the marker.

By progressively generating and refining a hypothesis, based on data derived from increasingly well-developed studies, tumor markers with clinical utility can be identified. Implementation of these recommendations when designing tumor marker studies will result in the generation and publication of appropriate and complete clinical data, leading to the adoption of new, well-validated tumor markers for routine clinical use.

DISCLOSURE

All the authors declared no competing interest.

REFERENCES

1. Sackett DL, Rosenberg WMC, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. BMJ 1996;312:71-72.

2. Price CP. Evidence-based laboratory medicine: supporting decision-making. Clin Chem 2000;46:1041-1050.

3. Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J. Potential benefits, limitations, and harms of clinical guidelines. BMJ 1999;318:527–30. http://bmj.com/cgi/content/full/318/7182/527/DC1 (Accessed May 2002).

4. Bast RC Jr, Ravdin P, Hayes DF et al. 2000 update of recommendations for the use of tumor markers in breast and colorectal cancer: clinical practice guidelines of the American Society of Clinical Oncology. J Clin Oncol 2001;19:1865–1878.

5. Hammond ME, Fitzgibbons PL, Compton CC et al. College of American Pathologists Conference XXXV: solid tumor prognostic factors-which, how and so what? Summary document and recommendations for implementation. Arch Pathol Lab Med 2000;124:958–965.

6. Clinical practice guidelines for the use of tumor markers in breast and colorectal cancer. Adopted on May 17, 1996 by the American Society of Clinical Oncology. J Clin Oncol 1996;14:2843–2877.

7. Bast RC Jr, Ravdin P, Hayes DF et al. Errata: 2000 update of recommendations for the use of tumor markers in breast and colorectal cancer: clinical practice guidelines of the American Society of Clinical Oncology. J Clin Oncol 2001;19:4185–4188.

8. Cobleigh MA, Vogel CL, Tripathy D et al. Multinational study of the efficacy and safety of humanized anti-HER2 monoclonal antibody in women who have HER2-overexpressing metastatic breast cancer that has progressed after chemotherapy for metastatic disease. J Clin Oncol 1999;17:2639–2648.

9. Hayes DF, Bast RC, Desch CE et al. Tumor marker utility grading system: a framework to evaluate clinical utility of tumor markers. J Natl Cancer Inst 1996;88:1456–1466.

10. Stearns V, Yamauchi H, Hayes DF. Circulating tumor markers in breast cancer: accepted utilities and novel prospects. Breast Cancer Res Treat1998;52:239–259.

11. McGuire WL, Clark GM. Prognostic factors and treatment decisions in axillary-node-negative breast cancer. N Engl J Med 1992;326:1756–1761.

12. Gasparini G, Pozza F, Harris AL. Evaluating the potential usefulness of new prognostic and predictive indicators in node-negative breast cancer patients. J Natl Cancer Inst 1993;85:1206–1219.

13. Hayes DF, Trock B, Harris AL. Assessing the clinical impact of prognostic factors: when is "statistically significant" clinically useful? Breast Cancer Res Treat 1998;52:305–319.

14. Isaacs C, Stearns V, Hayes DF. New prognostic factors for breast cancer recurrence. Semin Oncol 2001;28:53–67.

15. Hayes DF. Do we need prognostic factors in nodal-negative breast cancer? Arbiter. Eur J Cancer 2000;36:302–306.

16. Early Breast Cancer Trialists' Collaborative Group (EBCTG). Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. Lancet 2005;365:1687–1717.

17. McGuire WL, Chamness GC, Costlow ME et al. Steroids and human breast cancer. J Steroid Biochem 1975;6:723–727.

18. Allegra JC, Lippman ME, Thompson EB et al. Estrogen receptor status: an important variable in predicting response to endocrine therapy in metastatic breast cancer. Eur J Cancer 1980;16:323–331.

19. Jacobs TW, Gown AM, Yaziji H et al. Specificity of HercepTest in determining HER-2/neu status of breast cancers using the United States Food and Drug Administration-approved scoring system. J Clin Oncol1999;17:1983–1987.

20. Harvey JM, Clark GM, Osborne CK et al. Estrogen receptor status by immunohistochemistry is superior to the ligand-binding assay for predicting response to adjuvant endocrine therapy in breast cancer. J Clin Oncol 1999;17:1474–1481.

21. Altman DG, Lausen B, Sauerbrei W et al. Dangers of using "optimal" cutpoints in the evaluation of prognostic factors. J Natl Cancer Inst 1994;86:829–835.

22. Cristofanilli M, Budd GT, Ellis MJ et al. Circulating tumor cells, disease progression, and survival in metastatic breast cancer. N Engl J Med 2004;351:781–791.

23. Gelber RD, Bonetti M, Castiglione-Gertsch M et al. Tailoring adjuvant treatments for the individual breast cancer patient. Breast 2003;12: 558–568.

24. Slamon DJ, Clark GM, Wong SG et al. Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. Science 1987;235:177–182.

25. Yamauchi H, Stearns V, Hayes DF. The role of c-erbB-2 as a predictive factor in breast cancer. Breast Cancer 2001;8:171–183.