Methodological requirements for valid tissue-based biomarker studies that can be used in clinical practice

Lawrence D True, MD [Professor of Pathology]
Dept of Pathology, Box 356100; 1959 NE Pacific St. University of Washington Medical Center, Seattle, WA

Lawrence D True: ltrue@u.washington.edu

Abstract

Paralleling the growth of ever more cost efficient methods to sequence the whole genome in minute fragments of tissue has been the identification of increasingly numerous molecular abnormalities in cancers - mutations, amplifications, insertions and deletions of genes and patterns of differential gene expression, i.e. overexpression of growth factors and under-expression of tumor suppressor genes. These can be translated into assays to be used in clinical decision making. In general terms the result of such an assay is subject to a large number of variables regarding the characteristics of the available sample, particularities of the used assay and the interpretation of the results. This review discusses the effects of these variables on assays of tissue-based biomarkers, classified by macromolecule - DNA, RNA (including micro RNA, messenger RNA, long noncoding RNA, protein and phosphoprotein). As presently the majority of clinically applicable biomarkers concern immunohistochemically detectable proteins the review focuses on this area. However, the principles outlined are mostly applicable to any other analyte. A variety of pre-analytical variables impacts on the results obtained, including analyte stability (which is different for different analytes i.e. DNA, RNA or protein), period of warm and of cold ischemia, fixation time, tissue processing, sample storage time and storage conditions. In addition, assay variables play an important role, including reagent specificity (notably but not uniquely an issue concerning antibodies used in immunohistochemistry), technical components of the assay, quantitation and assay interpretation. Finally, appropriateness of an assay for clinical application is an important issue. Reference is made to publicly available guidelines to improve on biomarker development in general and requirements for clinical use in particular. Strategic goals are formulated in order to improve on the quality of biomarker reporting, including issues of analyte quality, experimental detail, assay efficiency and precision and assay appropriateness.

Keywords

immunohistochemistry; preanalytical; tissue; variables; prognostic; predictive

---

Introduction

Paralleling the growth of ever more cost efficient methods to sequence the whole genome of pieces of tissue as small as 0.01 grams has been identification of increasingly numerous genetic abnormalities in cancers - mutations, amplifications, insertions, deletions, and patterns of differential expression of genes, i.e. overexpression of growth factors and underexpression of tumor suppressor genes. Knowledge of tumor-associated molecular abnormalities has stimulated the large scale screening and development of reagents (small drugs, antibody derivatives) that bind to the abnormal genes or gene products. Pharmaceutical agents that are specific for these molecular targets have greatly increased in number. Succeeding this growth industry, which produces drugs that target these molecular abnormalities and that have both greater tumor-specific effect than conventional cytotoxic agents with less severe and less frequent side effects, has been the impetus to develop tissue-based biomarkers. Generically defined, tissue biomarkers include any tissue-based feature, ranging from the histological grade of a cancer to point mutations. We shall focus on molecular biomarkers in this review.

Biomarkers can be categorized as either **prognostic** or **predictive. Prognostic** markers predict clinical outcome after × number of years. The number of years appropriate to evaluate the value of a biomarker should be proportionate to the natural history of the cancer. The fate of most patients with lung cancer will become apparent within two years, in contrast, the outcome of the average patient with prostate cancer won't be known with certainty for more than five years. **Predictive** markers predict response of a tumor to a drug that putatively targets a key functional molecule involved in the biology of that tumor. Given the high cost of many targeted therapies, many early phase clinical trials are requiring or, for markers that are not integral to a clinical trial, encouraging the identification and validation of predictive biomarkers of these novel agents. However, identifying the biomarker that can serve as a detector of abnormalities in a functionally critical step of the progression of the cancer can be challenging, especially if the molecular pathway contains many regulatory genes. The mandate to identify a biomarker that can readily be assessed is particularly high when the predictive biomarker is integral to the trial. An **integral** marker is essential for a clinical trial, i.e. is used to qualify a patient for a specific therapy.

The performance of biomarkers is of concern to the regulatory agencies, the drug manufacturers, the health care deliverers and to patients, to maximize benefit and minimize costand complications of therapy based on expression of the biomarker. However, the performance of many biomarkers has been relatively poor [1].

There is a major challenge with immunohistochemical data that is used for tissue-based biomarkers. Biopharmaceutical companies that use the biomedical literature to identify predictive biomarkers or biomarkers that are integral to early phase clinical trials with novel drugs rely upon the literature to identify candidate predictive biomarkers and molecular targets for therapy. However, no more than fifty percent of studies can be replicated. Consequently, costs to develop drugs are often wasteful, incurred by companies on drugs for which the evidence of functional efficacy is poor. Theme of this review is how the quality of biomarker studies can be improved.

As attention has been directed to biomarkers, the biomed community has become aware of the effect of factors independent of the biologic state of the tumor that can lead to erroneous results of tissue-based assays. These can be categorized as pre-analytical and analytical variables that affect the expression and expression levels of biomarkers. These effects, which do not reflect the biology of the tumor, but are artifacts of tissue handling and of the way that the assay is performed, can skew results of the assays from what the true biologic state of the tissue is.

This introductory section discusses the effects of these variables on assays of tissue-based biomarkers, classified by macromolecule - DNA, RNA (including micro RNA, messenger RNA, long noncoding RNA), protein, and phosphoprotein. Many studies of the effect of preanalytical and analytical variables on macromolecules in tissue have been published. A repository of manuscripts of these studies, which were published in peer-reviewed journals, is on the website of the former Office of Biospecimen and Biorepository Research of NCI. Hundreds of studies have analyzed the effect of pre-analytical and analytical variables. A web-accessible library of publications, the Biospecimen Research Database (https:// brd.nci.nih.gov/BRN/brnHome.seam), organizes the publications in a table organized by type of specimen (blood, serum, plasma, urine, saliva, and tissue) and type of assay for each macromolecule (DNA, RNA, protein, small molecule, morphology, cell count/volume, and peptide).

General approach of this review

This review is focused on the most common materials used in by pathology labs in clinical practice – human tissue obtained at surgery, fixed in neutral buffered formalin and embedded in paraffin (FFPE tissue). Although the main emphasis is on the method most commonly used in clinical labs to evaluate biomarkers in FFPE tissue - immunoperoxidase stains of protein analytes assessed visually by pathologists - comments will be provided about other analytes and methods of assessment of tissue-based biomarkers. Finally, examples are based on studies of prostate cancer, the cancer with which the author is most familiar.

Pre-analytical variables

Stability

The stability of a potential biomarker varies with the type of macromolecule [2].

- **DNA** Although relatively resistant to degradation, the informativeness of DNA can be affected by tissue handling. Formaldehyde fragments DNA, resulting in shorter, less informative fragments wherein mutations might be missed, or, conversely, be misidentified. A relevant recent finding is that DNA sequences differ in different somatic tissues, based on whole genome sequences [3]. This finding raises the prospect that even the sequence of tissue used as "normal control DNA" may be misleading when used as reference DNA for the DNA sequence of a tumor in that individual.

- **mRNA** This most labile class of molecule has been used to generate molecular signatures of the effect of environment on tumor tissue. For example, RNA signatures of warm ischemia (the length of time devascularized tissue is at body temperature), cold ischemia (the length of time tissue sits at ambient temperature until stabilization by freezing or fixation) or even pre-operative diet (low vs. high protein content) have been generated. Micro RNA is a more stable nucleic acid than is messenger RNA.

- **Protein** One of the most widely used biomarker assays is immunohistochemical localization and assessment of expression level of specific proteins in FFPE tissue. At least sixty-two preanalytical variables that might affect immunohistochemical reactivity (IHR) have been evaluated [4]. These variables cover the time course from tissue acquisition through fixation and storage of the paraffin blocks to handling of sections prior to immunostaining. The extent that preanalytical conditions affect protein levels varies with the analyte and with the study. Below is a summary of those variables that are most frequently encountered, that contribute most to preanalytical variability, and that should be considered in immunohistochemical assays.

Delay in acquiring the tissue from the surgery

Laparoscopic procedures have resulted in longer warm ischemia time (the time between tissue devascularization and removal of tissue from the patient) than in the pre-laparoscopic era. Though changes in RNA profile are measurable, few studies have assessed the effect of warm ischemia on proteins. An evolving consensus is that warm ischemia time not exceed 2 hours.

Delay in fixation

In general, fixation can be delayed up to 12 hours without compromising IHR. In one study there was no effect on estrogen receptor (ER) or progesterone receptor (PR) IHR with delay in fixation of up to 8 hours at 19 °C or up to 12 hours at 4°C. Beyond 6 hours the effect of "cold" ischemia varies with the analyte.

Time in fixative

Fixation for at least 24 hours appears to be sufficient for preservation of all analytes. Though loss of immunoreactivity by underfixation is the most frequent concern, overfixation can affect staining of some analytes. For example, there is loss of signal when fixation of ER exceeds 4 days.

Post-fixation tissue processing

Although publications are discrepant, immunoreactivity can be variably affected, in an analyte-independent manner, at the following stages of tissue and section handling: duration and temperature of dehydration of fixed tissue, reagents used for clearing tissue and the temperature of clearing, duration and temperature of paraffin embedding, and the temperature and duration of heating to enhance adherence of sections to slides.

Duration of block storage

The length of time that FFPE blocks can be stored without loss of immunoreactivity varies by analyte and study. Some blocks have been stored as long as 25 years without loss of IHR of ER. IHR of some other analytes decreases after 2 years of block storage.

Duration and temperature of storage of sections

Preanalytical variables include paraffin coating of sections, desiccation, storage in nitrogen, and storage at −80 °C. The extent of effect varies with the study and the analyte.

Other variables—Labs are unlikely to change or control for some conditions despite weak evidence that the variable may affect IHR. These include the nature and pH of fixative (recommended: 10% neutral formalin buffered to pH 5–7), tissue:fixative ratio (recommended: at least 1:10), and size of tissue specimen (recommended: one dimension is no greater than 0.3 cm). The effect on IHR of other variables is minimal or negligible, such as type of tissue processor. And the effect of some variables has not been assessed, i.e. storage of specimens in different fixatives. Publications concur on the effect of fixation delay, fixation, dehydration, paraffin impregnation, drying of slides and storage conditions on immunostains, commenting that some recommendations by oncology interest groups differed from published guidelines. Fox example delay in fixation, recommended not to exceed one hour, can exceed 12 hours without discernible effect. And, sections mounted on slides should probably be stained within 6 days of preparation, though published guidelines cite maximum times of one and six weeks, respectively.

These studies have limitations. The methods used by the different investigators were not standardized. And a majority of studies have not been validated by independent investigations. What is needed are metrics to assess the state of retention of the in vivo state of proteins in tissue, prior to the effect on these analytes of preanalytical conditions. An initial set of markers that provide correlation, albeit weak, with quality of ER have been published. Loss of phosphorylated tyrosine immunoreactivity and increased acetylated lysine and HIF1A immunoreactivity correlate with decreased ER staining [5]. Unknown is whether these markers, or any set of markers, can be used to assess the degree of retention of the native state of all analytes. One suspects not, hypothesizing that host factors (variance in the functionality and metabolism of proteolytic enzymes) and analyte factors (primary and secondary structure, access to proteases), influence the quality of ischemic proteins.

Analytical variables

Stages in assaying proteins of interest are additional sources of ex-vivo variability.

- **Primary antibody**. Performance characteristics of the primary antibody include sensitivity and specificity for detecting the analyte of interest in sections of FFPE tissue. Rigorous assessment of specificity involves biochemical and cell biological techniques that clinical labs are virtually never prepared to undertake. These techniques include confirmation that engineered overexpression or knockdown of the analyte in cultures and preparations of cells (such as cell line microarrays) correlates with IHR. There are suggested guidelines for assessing the quality of

primary antibodies for use in IHC studies [6]. However, since most clinical labs are not prepared for detailed evaluations of the performance of antibodies, directors of IHC labs must rely on information provided by the antibody vendors. Too often this information lacks data relevant to IHC studies of FFPE tissue. For example, antibodies that produce a single band on Western gels and are localized to cells of interest in frozen tissue may not immunoreact with cells of interest in FFPE tissue.

- **Technical component of assays**. As currently practiced, immunoperoxidase stains involve at least 4 sequential chemical reactions. In principle conditions should be optimized for each step, i.e. antibodies should saturate binding sites and the kinetics of peroxidase catalysis of substrate diaminobenzidine should be optimized for the variables that affect the amount of substrate that is deposited on the tissue section – temperature, pH, concentration of electron donor hydrogen peroxide, duration of reaction. Such systematic optimization of all steps of IHC has been exemplified by a study correlating a biochemically assayed protein extracts from pituitary with immunoperoxidase stains quantified by densitometry [7]. Such optimization of the chemistry is virtually never done. Assays of protein in tissue that involve fewer reactions promise to be more reproducibile. Quantum dot-conjugated immunostains are, in principle, subject to less variance since there is no enzyme substrate reaction and fewer steps are needed [8]. Mass spectrometry analysis of protein content in cells of interest, i.e. cancer cells, is another method that is more quantitative than immunoperoxidase stains. However, there is currently no efficient way to isolate cells of interest from tissue in a volume sufficient for mass spectrometry.

- **Interpretive component of assays**. Currently immunoperoxidase stains are analyzed either visually by pathologists or quantified by image analysis. The advantage of visual interpretation is ready accessibility of stains without needing to acquire and devote effort using an image analysis system. However, visual cognition of the human eye varies with the task. While we can readily identify and categorize many objects that contrast with their microenvironment, such as cancer cells in tissue, human vision performs poorly in assessing degrees of difference of elements of an image, such as grey levels (http://bcs.mit.edu/people/adelson.html). Furthermore, tissue elements adjacent to the objects being analyzed influence our visual cognition [9]. Finally, there are biologic limitations to our ability to visually "measure" light intensity. For example, the sensitivity of our retina to light differs by the wavelength of light [10].

- **Quantitation**. In contrast to visual assessment, the optical density of the pixels that comprise the digitized image of an immunostain can be determined with great accuracy using an image analysis system. However, two components of image analysis limit transferability and reproducibility of the data. Both components involve a pathologist (or, plausibly, a pathology-competent individual). First, the cells of interest need to be identified and distinguished from histological mimics. No current image analysis system can reliably distinguish low-grade prostate carcinoma from such mimics as adenosis, basal cell hyperplasia and prostatic intraepithelial neoplasia. Second, a threshold must be set to distinguish pixels that are deemed "positive" (reaction product) from those that are "negative."

Interpathologist variance in these cognitive tasks is virtually never assessed. There are additional sources of variance in image analysis systems that should be considered, and either controlled for or discarded, as appropriate [11]. Finally, imaging systems need to include in their evaluation cells of interest that are stained below the positive threshold level. By using multiple stains (to identify nuclei, which would have to be annotated by the pathologist as cancer nuclei, and the analyte of interest) imaging systems can address many of these challenges [12].

- **Visual scoring.** How immunostain interpretations are reduced to single values that can be used by clinicians in making treatment decisions is an additional source of variance. Immunostains are characterized by two parameters – the percentage of cells that express analytes at different levels of intensity. The challenge has been incorporating the two parameters of immunostains into a single number. One frequently used calculation is the "HScore" (a sum of the products of level of stain intensity and percent of cells stained at each intensity level). A consequence of using the HScore is that stains that differ markedly can have the same value. A different way to report the data is as categories of staining [13]. Though categorical values lack the pseudo precision of HScores, their reproducibility is, in principle, higher.

Guidelines and Specifications

The literature can be made more robust by adhering to a number of specifications, such as reporting all of the components of immunohistochemical stains. A specification termed MISFISHIE has been published. This specification includes detailed characterization of primary antibodies used in IHC studies. However, the performance, specificity, and sensitivity of primary antibodies is not part of the MISFISHIE specification. Publications have made the point that specificity and sensitivity of primary antibodies should be validated. With respect to the quality of published IHC studies, and, thus, the probability that the results can be replicated, one cannot but think that the specification for information on preanalytical variables should be considered a part of the publication. Based upon reviews of IHC studies, publications should consider addressing the effect (if any) of warm and cold ischemia time, length of fixation, and tumor necrosis. These parameters could be reported as an addendum or as an incorporated table. This table could also be used by the editorial staff of the journals publishing IHC studies. A second reason for the inability to replicate IHC studies has to do with the interpretation of immunostains. Little attention, in general, has been paid to how they are interpreted. Two major sources of variance in interpretation are the fact that majority are accessed visually rather than by optical density of pixels in an image digitized file. Visual cognition is sufficiently poor that one questions whether no more than two, perhaps three, categories of stain intensity should be used. A second variable is that there are virtually never a set of standard curves, referencing the intensity (or, luminosity, when fluorochromes are used as the detection agents) to reference expression levels of the analytes. A third variable has to do with heterogeneity of tumors and a question of whether the samples chosen for the immunohistochemical analysis truly are representative of the state of expression of analytes in tumors. For example, the sequences of DNA from different regions of a large renal cell carcinoma differ [14]. The challenge is to

identify the region of the tumor that has genomic changes that correlate with tumor progression, i.e. growth rate, patterns of metastasis, resistance to therapy.

Predictive biomarkers, used to select patients for targeted therapy, will play an increasingly prominent role in cancer treatment as more drugs targeted to pathway molecules are made available. Several additional considerations pertain to these biomarkers – notably assay performance and validation, assay utility, and clinical utility.

Guidelines for tissue-based assays

A number of guidelines have been published within the past decade. MISFISHIE provides a checklist for components of tissue-based assays that are sufficiently detailed for an independent to be likely to successfully replicate the study [15]. REMARK criteria list the expected components of a clinical trial that has a tissue biomarker component [16]. Recently published is a set of guidelines for RNA array-based data that might be used as prognostic or predictive biomarkers [17]. Despite availability of these guidelines, there has been little incentive for most of the biomedical community to use them, presumably since adherence to them requires more effort and expense, both by the investigators who study the biomarkers and by the reviewers of manuscripts that report biomarker studies.

Use of biomarkers in practice

To effectively incorporate predictive biomarkers into practice requires close interaction between clinical trialists/oncologists, laboratorians/pathologists, outcome statisticians and regulatory agencies [18]. Appreciation has grown that the level of evidence of the predictive power of a biomarker provides a basis for assessing the value of the marker in clinical practice. Four categories of level of evidence have been proposed in studies of breast cancer biomarkers [19]. Levels of evidence range from anecdotal observations to validation of the performance of the biomarker in an independent cohort where clinicians and pathologists are blinded to the clinical outcome of patients in the study. The concept of level of evidence can be expanded to include evidence of clinical utility and validity in addition to the strength of evidence that the biomarker is predictive [20]. Stated another way, a predictive biomarker is of value if it categorizes a significant number of patients. Conversely a biomarker is of little value if most patients do or do not respond to the targeted treatment.

Once the performance of an assay has been characterized and deemed adequate for the intended purpose it should be validated. Validation of assays poses challenges for clinical labs, which may not have the specimens, resources, or experience to efficiently validate a tissue-based assay [21]. An additional consideration for immunohistochemical assays is the lack of a standard method for validating quantitative immunostains. In lieu of being able to generate standard deviations interobserver variance in quantitative values is often reported as kappa values, which is a non-parametric value [21].

A final step in accepting an assay for clinical use is approval by regulatory agencies, such as the FDA.

Future Directions (A wish list)

A challenge that pathologists are increasingly confronted with is assay reproducibility. Publications in which tissue-based assays provide clinical outcome data provide potentially valuable leads for investigators who are developing and testing the utility of drugs targeting gene products in specific molecular pathways. However, at least 50% of publications reporting the types of preclinical studies on which tissue-based biomarker assays are based are not reproducible [22]. Explanations are multifold, including insufficient detail in the original study, poor study design (such as lack of separate training and independent test sets), and lack of incentives to robustly validate the data. Biomarker studies could presumably be improved based on the following wish list:

- **Details of Experiments**. That publications that include tissue-based biomarkers contain all methodological information of the biomarker assay sufficient for an independent investigator to conduct and interpret the results of the assay identical to how the assay was done in the original publication [23]. As discussed above, such a specification, with spreadsheet, which incorporates most components of tissue-based assays in a checklist form (MISFISHIE) has been published [15].

- **Efficiency**. That assays be done and interpreted in the most effort-efficient manner possible for that assay.

- **Precision of Tissue-based Assays**. That variability of interpretation of the assay be reported in all publications. This goal involves more than one investigator independently interpreting the assay. Although the type of assay with presumably the greatest inter-observer variance is an immunoperoxidase assay that is visually interpreted, assays that use image analysis should be included since pathologists make several usually subjective decisions, as previously discussed:

  - Selection of the region of interest that is digitized and analyzed.

  - Setting a threshold for binarizing results.

  - A method that evaluates objects that are "negative", i.e. that stain at or below the level of background staining.

Analogous to solution assays, the results of each pathologist could be reported as a range of values, which reflects the precision of the assay.

That levels of assay precision for clinical management of the patient be determined and reported for all assays.

- **Accuracy of Tissue-based Assays**. That a set of standards be developed that includes the dynamic range of the assay and the range of the majority of values of the type of cancer for which the assay is used.

- **Analyte quality**. That internal standards that assess the quality of protein in tissue be developed.

- **Assay appropriateness**. That results of the assay be "fit for purpose." Included in this wish is the expectation that the range of inter- and intra-observer or test variance of the assay that is acceptable to the clinician be reported.

Acknowledgments

References

1. Meshinchi S, Hunger SP, Aplenc R, Adamson PC, Jessup JM. Lessons learned from the investigational device exemption review of Children's Oncology Group trial AAML1031. Clin Cancer Res. 2012; 18:1547–1554. [PubMed: 22422407]

2. Hewitt SM, Badve SS, True LD. Impact of preanalytic factors on the design and application of integral biomarkers for directing patient therapy. Clin Cancer Res. 2012; 18:1524–1530. [PubMed: 22422404]

3. O'Huallachain M, Karczewski KJ, Weissman SM, Urban AE, Snyder MP. Extensive genetic variation in somatic human tissues. Proc Natl Acad Sci U S A. 2012; 109:18018–18023. [PubMed: 23043118]

4. Engel KB, Moore HM. Effects of preanalytical variables on the detection of proteins by immunohistochemistry in formalin-fixed, paraffin-embedded tissue. Arch Pathol Lab Med. 2011; 135:537–543. [PubMed: 21526952]

5. Neumeister VM, Anagnostou V, Siddiqui S, England AM, Zarrella ER, Vassilakopoulou M, Parisi F, Kluger Y, Hicks DG, Rimm DL. uantitative assessment of effect of preanalytic cold ischemic time on protein expression in breast cancer tissues. J Natl Cancer Inst. 2012; 104:1815–1824. [PubMed: 23090068]

6. Bordeaux J, Welsh A, Agarwal S, Killiam E, Baquero M, Hanna J, Anagnostou V, Rimm D. Antibody validation. Biotechniques. 2010; 48:197–209. [PubMed: 20359301]

7. Gross DS, Rothfeld JM. Quantitative immunocytochemistry of hypothalamic and pituitary hormones: validation of an automated, computerized image analysis system. J Histochem Cytochem. 1985; 33:11–20. [PubMed: 2578140]

8. Zrazhevskiy P, Gao X. Multifunctional Quantum Dots for Personalized Medicine. Nano Today. 2009; 4:414–428. [PubMed: 20161004]

9. Fandel TM, Pfnür M, Schäfer SC, Bacchetti P, Mast FW, Corinth C, Ansorge M, Melchior SW, Thüroff JW, Kirkpatrick CJ, Lehr HA. Do we truly see what we think we see? The role of cognitive bias in pathological interpretation. J Pathol. 2008; 216:193–200. [PubMed: 18683853]

10. Wells WA, Rainer RO, Memoli VA. Equipment, standardization, and applications of image processing. J Clin Pathol. 1993; 99:48–56.

11. True LD. Quantitative immunohistochemistry: a new tool for surgical pathology? Am J Clin Pathol. 1988; 90:324–325. [PubMed: 3046322]

12. Rubin MA, Zerkowski MP, Camp RL, Kuefer R, Hofer MD, Chinnaiyan AM, Rimm DL. Quantitative determination of expression of the prostate cancer protein alpha-methylacyl-CoA racemase using automated quantitative analysis (AQUA): a novel paradigm for automated and continuous biomarker measurements. Am J Pathol. 2004; 164:831–840. [PubMed: 14982837]

13. Etzioni R, Hawley S, Billheimer D, True LD, Knudsen B. Analyzing patterns of staining in immunohistochemical studies: application to a study of prostate cancer recurrence. Cancer Epidemiol Biomarkers Prev. 2005; 14:1040–1046. [PubMed: 15894650]

14. Gerlinger M, Rowan AJ, Horswell S, Larkin J, Endesfelder D, Gronroos E, Martinez P, Matthews N, Stewart A, Tarpey P, Varela I, Phillimore B, Begum S, McDonald NQ, Butler A, Jones D, Raine K, Latimer C, Santos CR, Nohadani M, Eklund AC, Spencer-Dene B, Clark G, Pickering L, Stamp G, Gore M, Szallasi Z, Downward J, Futreal PA, Swanton C. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N Engl J Med. 2012; 366:883–892. [PubMed: 22397650]

15. Deutsch EW, Ball CA, Berman JJ, Bova GS, Brazma A, Bumgarner RE, Campbell D, Causton HC, Christiansen JH, Daian F, Dauga D, Davidson DR, Gimenez G, Goo YA, Grimmond S, Henrich T, Herrmann BG, Johnson MH, Korb M, Mills JC, Oudes AJ, Parkinson HE, Pascal LE, Pollet N, Quackenbush J, Ramialison M, Ringwald M, Salgado D, Sansone SA, Sherlock G, Stoeckert CJ Jr, Swedlow J, Taylor RC, Walashek L, Warford A, Wilkinson DG, Zhou Y, Zon LI, Liu AY, True LD. Minimum information specification for in situ hybridization and immunohistochemistry experiments (MISFISHIE). Nat Biotechnol. 2008; 26:305–312. [PubMed: 18327244]

16. McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM. Statistics Subcommittee of the NCI-EORTC Working Group on Cancer Diagnostics. Reporting recommendations for tumor marker prognostic studies (REMARK). J Natl Cancer Inst. 2005; 97:1180–1184. [PubMed: 16106022]

17. McShane LM, Cavenagh MM, Lively TG, Eberhard DA, Bigbee WL, Williams PM, Mesirov JP, Polley MY, Kim KY, Tricoli JV, Taylor JM, Shuman DJ, Simon RM, Doroshow JH, Conley BA. Criteria for the use of omics-based predictors in clinical trials. Nature. 2013; 502:317–320. [PubMed: 24132288]

18. Poste G, Carbone DP, Parkinson DR, Verweij J, Hewitt SM, Jessup JM. eveling the playing field: bringing development of biomarkers and molecular diagnostics up to the standards for drug development. J Natl Cancer Inst. 1996; 88:1456–1466. [PubMed: 8841020]

19. Hayes DF, Bast RC, Desch CE, Fritsche H Jr, Kemeny NE, Jessup JM, Locker GY, Macdonald JS, Mennel RG, Norton L, Ravdin P, Taube S, Winn RJ. Tumor marker utility grading system: framework to evaluate clinical utility of tumor markers. Clin Cancer Res. 2012; 18:1515–1523. [PubMed: 22422403]

20. Febbo PG, Ladanyi M, Aldape KD, De Marzo AM, Hammond ME, Hayes DF, Iafrate AJ, Kelley RK, Marcucci G, Ogino S, Pao W, Sgroi DC, Birkeland ML. NCCN Task Force report: Evaluating the clinical utility of tumor markers in oncology. J Natl Compr Canc Netw. 2011; 9(Suppl 5):S1–S32.

21. Williams PM, Lively TG, Jessup JM, Conley BA. Bridging the gap: moving predictive and prognostic assays from research to clinical use. Clin Cancer Res. 2012; 18:1531–1539. [PubMed: 22422405]

22. Begley CG, Ellis LM. Drug development: Raise standards for preclinical cancer research. Nature. 2012; 483:531–533. [PubMed: 22460880]

23. Leong AS, Leong TY. Standardization in immunohistology. Methods Mol Biol. 2011; 724:37–68. [PubMed: 21370005]