NCCN Task Force Report: Evaluating the Clinical Utility of Tumor Markers in Oncology

Phillip G. Febbo, MD; Marc Ladanyi, MD; Kenneth D. Aldape, MD; Angelo M. De Marzo, MD, PhD; M. Elizabeth Hammond, MD; Daniel F. Hayes, MD; A. John Iafrate, MD, PhD; R. Kate Kelley, MD; Guido Marcucci, MD; Shuji Ogino, MD, PhD, MS; William Pao, MD, PhD; Dennis C. Sgroi, MD; and Marian L. Birkeland, PhD

The National Comprehensive Cancer Network® (NCCN®) appreciates that supporting companies recognize NCCN's need for autonomy in the development of the content of NCCN resources. All NCCN content is produced completely independently. The NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines®) are not intended to promote any specific therapeutic modality. The distribution of this task force report is supported by educational grants from Dendreon Corporation, Genomic Health, Inc., Genoptix Medical Laboratory, and Pfizer.
Postal and Contact Information

JNCCN (ISSN 1540-1405, USPS Publication Number 240), the official journal of the National Comprehensive Cancer Network® (NCCN®), is published monthly by Harborside Press®, 37 Main Street, Cold Spring Harbor, NY 11724. Periodicals postage paid at Cold Spring Harbor, NY and additional mailing offices.

Change of Address: Postmaster: send address changes to JNCCN, c/o Harborside Press®, 37 Main Street, Cold Spring Harbor, NY 11724. Recipient: to change your address contact subscriptions@harborsidepress.com or fax 631-692-0905. Please state that this change of address request is for JNCCN.

© National Comprehensive Cancer Network, Inc. 2011. All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means now or hereafter known, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from NCCN.

Subscriptions: Prices for yearly subscriptions (12 issues plus supplements) are: Individual: Print only or online only, US $495; Can/Mex $595; Int’l $610; print and online, US $545; Can/Mex $670; Int’l $690. Institutional: Print only or online only, US $790; Can/Mex $893; Int’l $910; print and online, US $860; Can/Mex $960; Int’l $1035. Single Copy: US $90; Can/Mex $110; Int’l $125.00. Subscription Inquiries should be directed to Wendy McGullam, Harborside Press®, at: 631-935-7651 or wendy@harborsidepress.com. Online access is available to subscribers through HighWire Press (JNCCN.org).

Contact Information

Editorial Office: Manuscripts, correspondence, and manuscripts for publication should be sent to Kimberly Callan, Senior Director, Professional and Patient Publications, JNCCN, 275 Commerce Drive, Suite 300, Fort Washington, PA 19034; or e-mail JNCCN@nccn.org. Correspondence can also be faxed: 215-690-0283 (attn: JNCCN). Questions about requirements for publication or topic suitability can be directed as above or to Harold J. Burstein, MD, PhD, Editor-in-Chief, JNCCN, 275 Commerce Drive, Suite 300, Fort Washington, PA 19034; or e-mail burstein@nccn.org.

For instructions for authors published in JNCCN as space allows and can be found online at www.JNCCN.org. They can also be requested by calling 215-690-0270 or e-mailing JNCCN@nccn.org.

Advertising

To purchase advertising space: Contact David Horowitz, Director of Business Development, Harborside Press®, 37 Main Street, Cold Spring Harbor, NY 11724; phone 631-935-7652; fax 631-692-0805; or e-mail david@harborsidepress.com.

To send ad materials: Ship to Harborside Press®, Attn: Wendy McGullam, (JNCCN, Vol __ Issue __), 37 Main Street, Cold Spring Harbor, NY 11724; phone 631-935-7651; fax 631-692-0805; or e-mail wendy@harborsidepress.com.


Production

Reprints: Reprints of individual articles are available. Orders must be for a minimum of 100 copies. Please contact David Horowitz, Director of Business Development, Harborside Press®, 37 Main Street, Cold Spring Harbor, NY 11724; phone 631-935-7652; fax 631-692-0805; or e-mail david@harborsidepress.com.

Permissions

For information about photocopying, republishing, reprinting, or adapting material, please go online to www.NCCN.org/about/permissions.

Indexing

JNCCN is indexed by MEDLINE/PUBMED®, Chemical Abstracts, EMBASE, EmCare, and Scopus. This paper meets the requirements of ANSI/NISO Z39.48-1992 (Permanence of Paper) effective with Volume 1, Issue 1, 2003.

JNCCN is a member of the Medscape Publisher’s Circle®, an alliance of leading medical publishers whose content is featured on Medscape (http://www.medscape.com). Medscape is part of the WebMD Medscape Health Network, a leading online healthcare resource for professionals and consumers.

Disclaimer

The treatment algorithms presented in JNCCN and its supplements are a statement of evidence and consensus of the authors regarding their views of currently accepted approaches to treatment. Any clinician seeking to apply or consult these guidelines is expected to use independent medical judgment in the context of individual circumstances to determine any patient’s care or treatment. The research articles, reviews, and other individually authored papers presented herein are the work of the authors listed. Furthermore, the reader is advised that, except where specifically stated, all of the ideas and opinions expressed in JNCCN are the authors’ own and do not necessarily reflect those of NCCN, the member organizations, the editor, or the publisher. Publication of an advertisement or other product mention in JNCCN should not be construed as an endorsement of the product or the manufacturer’s claims.

The information contained in JNCCN is presented for the purpose of educating our readership on cancer treatment and management. The information should not be relied on as complete or accurate, nor should it be relied on to suggest a course of treatment for a particular individual. It should not be used in place of a visit, call, consultation, or the advice of a licensed physician or other qualified health care provider. Patients with health care-related questions or concerns are advised to contact a physician or other qualified health care provider promptly.

Although every attempt has been made to verify that information presented within is complete and accurate, the information is provided “AS IS” without warranty, express or implied. NCCN hereby excludes all implied warranties of merchantability and fitness for a particular use or purpose with respect to the Information. Furthermore, NCCN makes no warranty as to the reliability, accuracy, timeliness, usefulness, adequacy, completeness, or suitability of the information.
JNCCN is dedicated to improving the quality of cancer care locally, nationally, and internationally while enhancing the collaboration between academic medicine and the community physician. JNCCN is further committed to disseminating information across the cancer care continuum by publishing clinical practice guidelines and reporting rigorous outcomes data collected and analyzed by experts from the world’s leading care centers. JNCCN also provides a forum for original research and review papers focusing on clinical and translational research and applications of the NCCN Guidelines in everyday practice, as well as correspondence and commentary.

Mission Statement

The National Comprehensive Cancer Network®, a not-for-profit alliance of 21 of the world’s leading cancer centers, is dedicated to improving the quality and effectiveness of care provided to patients with cancer. Through the leadership and expertise of clinical professionals at NCCN Member Institutions, NCCN develops resources that present valuable information to the numerous stakeholders in the health care delivery system. As the arbiter of high-quality cancer care, NCCN promotes the importance of continuous quality improvement and recognizes the significance of creating clinical practice guidelines appropriate for use by patients, clinicians, and other health care decision-makers. The primary goal of all NCCN initiatives is to improve the quality, effectiveness, and efficiency of oncology practice so patients can live better lives. For more information, visit www.NCCN.org.
CME Accreditation

The National Comprehensive Cancer Network (NCCN) is accredited by the Accreditation Council for Continuing Medical Education (ACCME) to provide continuing medical education for physicians. The NCCN designates this journal-based CME activity for a maximum of 0.75 AMA PRA Category 1 Credits™. Physicians should only claim credit commensurate with the extent of their participation on the activity.

This educational activity was planned and produced in accordance with ACCME Essential Areas and Policies.

The NCCN adheres to the ACCME Standards for Commercial Support of Continuing Medical Education.

The NCCN is accredited as a provider of continuing nursing education by the American Nurses Credentialing Center’s (ANCC) Commission on Accreditation.

This activity is approved for 0.75 contact hours. Approval as a provider refers to recognition of educational activities only and does not imply ANCC Commission on Accreditation approval or endorsement of any product.

Kristina M. Gregory, RN, MSN, OCN, is our nurse planner for this educational activity.

Target Audience

This educational program is designed to meet the needs of oncologists, advanced practice nurses, and other clinical professionals who treat and manage patients with cancer.

Educational Objectives

After completion of this CME/CE activity, participants should be able to:

• Distinguish between a prognostic factor and a predictive factor as they apply to oncology
• Define the term “clinical utility” as it applies to the incorporation of tumor markers into clinical practice guidelines
• Describe the factors considered in the evaluation of the clinical utility of a tumor marker in oncology
• Describe the complementary roles played by pathologists and clinicians in ensuring accurate testing of tumor markers
• Explain the difference between the terms “laboratory certification” and “laboratory accreditation” as they apply to pathology laboratories performing tumor marker testing
• Describe the importance of pathologic expertise, internal quality control standards, and external quality assurance monitoring in determining the reliability of tumor marker testing
• Explain the significance of FDA approval as it applies to a diagnostic test
• Summarize initiatives for transitioning basic and translational research findings on tumor markers into clinical practice

The opinions expressed in this publication are those of the participating faculty and not those of NCCN, Dendreon Corporation, Genomic Health, Inc., Genoptix Medical Laboratory, Pfizer, or the manufacturers of any products mentioned herein.

This publication may include the discussion of products for indications not approved by the FDA.

Participants are encouraged to consult the package inserts for updated information and changes regarding indications, dosages, and contraindications. This recommendation is particularly important with new or infrequently used products.

Activity Instructions

Participants will read all portions of this monograph, including all tables, figures, and references. To receive your continuing education credit and certificate, visit http://www.cvent.com/d/rcq8rk to complete the post-test and evaluation. A minimum passing score of 70% is required on the post-test to be eligible for credits. If a minimum score is not achieved, you will be sent an e-mail with the opportunity to retake the test.

All post-test scores must be received by midnight on November 25, 2012 in order to be eligible for credit.

It should take approximately 0.75 hours (45 minutes) to complete the activity as designed. There is no registration fee for this activity. Certificates for passing scores will be e-mailed within 15 business days of submission of post-test and evaluation.

Copyright 2011, National Comprehensive Cancer Network (NCCN). All rights reserved. No part of this publication may be reproduced or transmitted in any other form or by any means, electronic or mechanical, without first obtaining written permission from the NCCN.
## Disclosure of Affiliations and Significant Relationships

Dr. Febbo has disclosed that he has no financial interests, arrangements, or affiliations with the manufacturers of any products or devices discussed in this report or their competitors.

Dr. Ladanyi has disclosed that he has financial interests, arrangements, or affiliations with the manufacturers of any products or devices discussed in this report or their competitors. He is a speaker for Infinity Pharmaceuticals.

Dr. Aldape has disclosed that he has no financial interests, arrangements, or affiliations with the manufacturers of any products or devices discussed in this report or their competitors.

Dr. Becker has disclosed that he has no financial interests, arrangements, or affiliations with the manufacturers of any products or devices discussed in this report or their competitors.

Dr. Birkeland has disclosed that she has financial interests, arrangements, or affiliations with the manufacturers of any products or devices discussed in this report or their competitors. She owns stock in GlaxoSmithKline. She is an employee of the National Comprehensive Cancer Network.

Dr. Conley has disclosed that she has financial interests, arrangements, or affiliations with the manufacturers of any products or devices discussed in this report or their competitors. She received clinical research support from sanofi-aventis; Endocyte, Inc.; and ImClone Systems.

Dr. De Marzo has disclosed that he has no financial interests, arrangements, or affiliations with the manufacturers of any products or devices discussed in this report or their competitors.

Dr. Hammond has disclosed that she has no financial interests, arrangements, or affiliations with the manufacturers of any products or devices discussed in this report or their competitors.

Dr. Hayes has disclosed that he has financial interests, arrangements, or affiliations with the manufacturers of any products or devices discussed in this report or their competitors. He receives clinical research support from GlaxoSmithKline; Pfizer Inc.; Novartis AG; and Veridex, LLC. He is a consultant for Chugai Pharmaceutical Co., Ltd. and BioMarker Strategies, and owns stock in Oncimmune LLC. He also holds patents with Immunicon Corporation and University of Michigan.

Dr. Iafrate has disclosed that he has financial interests, arrangements, or affiliations with the manufacturers of any products or devices discussed in this report or their competitors. He is a consultant for Clovis Oncology; Bristol-Myers Squibb Company; AstraZeneca Pharmaceuticals LP; MolecularMD; and Symphony Evolution, Inc. He received license compensation from MolecularMD.

Dr. Ogino has disclosed that he has no financial interests, arrangements, or affiliations with the manufacturers of any products or devices discussed in this report or their competitors.

Dr. Pao has disclosed that he has financial interests, arrangements, or affiliations with the manufacturers of any products or devices discussed in this report or their competitors. He receives clinical research support from Enzon Pharmaceuticals, Inc.; AstraZeneca Pharmaceuticals LP; MolecularMD; and Symphony Evolution, Inc. He receives license compensation from MolecularMD.

Dr. Sgroi has disclosed that he has financial interests, arrangements, or affiliations with the manufacturers of any products or devices discussed in this report or their competitors. He receives patent compensation from bioTheranostics, Inc.
NCCN Task Force Report: Evaluating the Clinical Utility of Tumor Markers in Oncology

Phillip G. Febbo, MD; Marc Ladanyi, MD; Kenneth D. Aldape, MD; Angelo M. De Marzo, MD, PhD; M. Elizabeth Hammond, MD; Daniel F. Hayes, MD; A. John Iafrate, MD, PhD; R. Kate Kelley, MD; Guido Marcucci, MD; Shuji Ogino, MD, PhD, MS; William Pao, MD, PhD; Dennis C. Sgroi, MD; and Marian L. Birkeland, PhD

Key Words
NCCN, tumor marker, biomarker, analytic validity, clinical validity, clinical utility, predictive marker, prognostic marker, companion diagnostic marker, molecular marker

Abstract
The molecular analysis of biomarkers in oncology is rapidly advancing, but the incorporation of new molecular tests into clinical practice will require a greater understanding of the genetic changes that drive malignancy, the assays used to measure the resulting phenotypes and genotypes, and the regulatory processes that new molecular biomarkers must face to be accepted for clinical use. To address these issues and provide an overview of current molecular testing in 6 major malignancies, including glioma, breast cancer, colon cancer, lung cancer, prostate cancer, and acute myelogenous leukemia, an NCCN Task Force was convened on the topic of evaluating the clinical utility of tumor markers in oncology. The output of this meeting, contained within this report, describes the ways biomarkers have been developed and used; defines common terminology, including prognostic, predictive, and companion diagnostic markers, and analytic validity, clinical validity, and clinical utility; and proposes the use of a combination level of evidence score to aid in the evaluation of novel biomarker tests as they arise. The current state of regulatory oversight and anticipated changes in the regulation of molecular testing are also addressed. (JNCCN 2011;9[Suppl 5]:S1–S32)

The science and practice of oncology are rapidly evolving based on more complete knowledge regarding cancer genomes and specific genetic events driving malignant disease. This knowledge, coupled with advances in medicinal chemistry, has ushered in an era of molecular oncology care in which, increasingly, best care practices require molecular insight. Although tumor markers previously played a role in a relatively small fraction of patients diagnosed with cancer, actionable molecular assays are now available that may guide treatment decisions for the most common, deadly malignancies. Given the rapidity with which the field is changing, NCCN assembled a task force to help educate the community with respect to the terminology of contemporary tumor markers and provide the current state of biomarker validation for the most common forms of cancer.

Although biomarker discovery is thriving, incorporation of biomarkers into clinical practice lags behind. Some of the challenges to clinical adoption include assay variability and inadequate analytic validation; poor study design and analysis; and inadequate reporting, and practical obstacles such as lack of resources, personnel, or expertise in smaller clinical laboratories. In addition, a growing number of multi-analyte tumor markers require comprehensive technologies and computational algorithms. These platforms and analytic approaches may be understandable to few practitioners and represent challenges to independent validation and verification. It is imperative that the field of oncology works with a common language and clear standards of evidence so that the merits of established and emerging tumor markers can be communicated in a clear and unambiguous manner.

This report provides the basic definitions for biomarkers in oncology, briefly reviews the steps necessary for the clinical development of a biomarker, and sum-
marizes a previously published approach to categorizing the strength of evidence supporting the use of a biomarker in a specific clinical situation. The task force hopes that this information will be useful to investigators who are discovering and/or validating new biomarkers and molecular tests, and to clinicians, guideline developers, and other stakeholders, such as health policymakers and payors, who participate in decision-making regarding the incorporation of novel biomarkers into clinical oncology. The final section of this report uses examples from 6 important disease areas to illustrate both the merits of tests that are currently in clinical use and the evidence that must be assembled for other tests to be accepted for general use. Through analyzing these examples of established and emerging biomarkers, the task force hopes to codify the appropriate evaluation of the required level of evidence necessary for making recommendations for incorporating new tests into clinical practice guidelines and general clinical use.

Types of Molecular Markers
Molecular markers are developed to address a variety of indications. An individual marker may serve more than one purpose and thus can fall into more than one category of biomarker. In addition, a single biomarker may have different categorization across tumor types and/or stages of disease. This section defines common terms used to describe molecular tests.

Diagnostic Markers
A large category of molecular tests aid in the diagnosis or subclassification of a particular disease state. Diagnostic subclassification may result in different management of the disease, but the marker is used primarily to establish the particular disease that is present in the patient sample. Examples of diagnostic molecular tests include immunophenotyping in non–Hodgkin’s lymphoma, and fluorescence in situ hybridization (FISH) to show the presence of the Philadelphia chromosome in chronic myelogenous leukemia.

Prognostic Markers
Prognostic markers have an association with some clinical outcomes, such as overall survival or recurrence-free survival, independent of the treatment rendered. An example of a prognostic marker is the presence of p53 mutations, which identify subsets of patients who will have a more aggressive disease course for certain cancers, regardless of current treatment options.

Predictive Markers
Predictive markers predict the activity of a specific class or type of therapy, and are used to help make more specific treatment decisions. They are used as indicators of the likely benefit of a specific treatment to a specific patient. Human epidermal growth factor receptor 2 (HER2[ERBB2]) is one example of a predictive marker. HER2[ERBB2]-negative tumors do not respond to trastuzumab, and therefore HER2[ERBB2]-positivity is predictive of potential trastuzumab response in a patient with newly diagnosed breast cancer.

Companion Diagnostic Markers
Companion diagnostic markers may be diagnostic, prognostic, or predictive, but are used to identify a subgroup of patients for whom a therapy has shown benefit. Specifically, prospective data show that patients with positive markers benefit from the therapy. Thus, although companion diagnostic markers are primarily a subgroup of predictive markers, evidence may not be sufficient to determine whether they have independent prognostic or predictive strength for the disease or class of therapy. One recent example of co-approval of a drug and companion diagnostic is the BRAF V600E mutation test coapproved with the kinase inhibitor vemurafenib. BRAF mutations are found in 30% to 60% of melanomas, and the kinase activating BRAF V600E mutation confers sensitivity to vemurafenib, a small molecule BRAF inhibitor. Package insert material for the drug specifies that a test must be performed to establish whether the tumor carries a BRAF V600E mutation.

Types of Validation
For a test to become generally useful, it must have demonstrated analytic validity and clinical utility. The former includes its reproducibility and quality as a test. Clinical validity implies that the marker identifies 2 groups that can be distinguished biologically and have different outcomes, but this observation may not indicate that it should be used for routine clinical care. Clinical utility implies that high-level evidence shows that use of the marker improves patient outcome sufficiently to justify its
incorporation into routine clinical care. This section reviews the terms analytic validation, clinical validation, and clinical utility.

**Analytic Validation**

Analytic validation focuses on determining how accurately and reliably the assay measures the molecular event of interest. Assays used in a research setting may be extremely reliable in the hands of skilled and meticulous laboratory investigators, but not scalable or feasible in a general clinical setting with routinely available tissue samples. Even assays that are routinely performed in the laboratory require analytic validation within a clinical laboratory setting to be used in making clinical decisions. Significant differences exist between research laboratories and clinical laboratories, and any assay being developed for the management of patients must be established within a clinical laboratory. For example, immunohistochemistry on formalin-fixed paraffin-embedded (FFPE) tumor samples is a well-established technique widely used in research and essential clinical cancer testing, such as in estrogen receptor 1 (ER-α[ESR1]) and progesterone receptor (PgR[PR]) or HER2 (ERBB2) staining in breast cancer. However, results can be influenced by a variety of factors, both within and outside the control of the pathology laboratory. One step that is often neglected is adequate validation of commercial or other antibodies. In one case, a commercial antibody presumably directed against telomerase was subsequently shown to recognize nucleolin.

To ensure reproducible findings, clinical laboratories should understand the impact that preanalytic variables and specimen processing have on assay performance. Simply stated, analytic validation ensures that the same answer will be produced for the same sample within predefined technical variation. Analytic validation does not signify that the result itself has any clinical relevance.

In recognition of the critical importance of analytic validation for biomarkers, multiple groups have developed recommendations and frameworks with which to standardize the assessment. One effort designed to provide a framework to record and potentially control some of the preanalytic specimen-handling variables is the Biospecimen Reporting for Improved Study Quality (BRISQ) recommendations, which resulted from a committee formed at the Biospecimen Research Network Symposium in 2009. Input was received from a wide range of experienced scientists, clinicians, pathologists, statisticians, and professionals from other disciplines, and the resulting publication can be used to assemble a checklist of elements that should be reported for clinical biospecimens, including information about the patient, diagnosis, tissue handling, fixation times, and storage and transport of tissue samples. Although not all information may be available or appropriate for each specimen type, use of a checklist and recommendations such as these can help to ensure that test results can be properly interpreted, whether the samples are used for immunohistochemistry, FISH, mutation testing, or multigene expression assays.

**Clinical Validation**

Clinical validation assesses the strength of association between the assay results and the clinical outcome of interest, whether it is diagnostic, prognostic, or predictive. A large number of measures are used to assess these associations. Although measures of statistical significance are important, robust measures of the strength of association, such as receiver operator characteristic (ROC) curves, the area under the curve of ROC analysis, and the sensitivity, specificity, and positive predictive value of assays, are critical to understanding the strength of their clinical association and help bridge analysis to clinical utility (discussed later). These analyses address whether one can be sure the clinical state is positive if the test is positive (positive predictive power), and vice versa (negative predictive power).

No single best design exists for validation studies, but some general observations can be made. First, biomarkers are often developed within a single dataset and resampling methods, such as leave-one-out cross-validation, fold-validation, or bootstrapping, are used to set assay parameters and estimate performance. Although these approaches help determine if a stable assay can be developed, they often underestimate the true error rate. For larger datasets, identifying a priori a “discovery” or “training” subset of samples and a “validation” or “test” subset of samples is common practice. The assay’s characteristics are tested and refined in the discovery subset and, when the assay parameters are fixed, applied to the validation subset. Although this minimizes bias and overfitting within the dataset, a concern remains that characteristics specific to the dataset being used may result in validation that does not generalize to
other sample or patient populations. Thus, independent validation sets are required for the most rigorous clinical validation of an assay.

Evaluation of new biomarkers can be aided if tumor biomarker studies and the journals reporting them adhere to the Reporting Recommendations for Tumor Marker Prognostic Studies (REMARK). These guidelines were produced by a workgroup convened as part of the NCI-EORTC joint meeting on Cancer Diagnostics in Nyborg, Denmark, in 2000. REMARK criteria include specification of patient populations, biological specimen under study, assay methods, study design, and statistical methods, and detailed guidelines for analysis and presentation of data. This reporting standard is now requested for manuscripts being submitted to journals such as the *Journal of Clinical Oncology*. Although the tissue resources are not always available to allow the most rigorous validation of assays, providing a standardized means to communicate the level of clinical validation is critical to biomarker development and underpins the motivations behind the levels of evidence discussed below. Using reporting standards such as these should help establish the clinical validity of new biomarkers being used in cancer research and treatment.

A good example of rigorous clinical validation is seen in a publication focused on validating several independently derived gene expression–based prognostic markers for early-stage lung cancer.21,22 In this study, 442 samples were collected from 6 institutions and processed at 1 of 4 laboratories using a single, reproducible protocol, and data were analyzed using a blinded validation step.21 Previous gene expression studies have been hampered by small sample size, samples having been collected from only a single institution, and lack of uniformity in sample processing and data analysis.

**Clinical Utility**

Clinical utility refers to the ability of the assay to improve clinical decision-making and patient outcomes. Clinical utility depends on the clinical situation, availability of effective therapies, magnitude of clinical benefit (or lack thereof) in one marker–designated group versus another, and relative value the patient, caregiver, and society place on the differences in benefits and risks in these separate groups, and their perceptions of these differences. For example, even if a marker clearly distinguishes outcomes between “positive” and “negative” groups, if evidence for differential treatment is not available, then no reason exists to test for the marker. Likewise, the marker may separate outcomes of 2 groups, but if the therapy is very effective in both, just slightly less effective in one, the marker does not have clinical utility.

An emerging standard for the adoption of new molecular tests is the demonstration of clinical utility. A novel assay may have outstanding analytic validity and proven clinical validity but may strongly correlate with an established clinical or histopathologic predictor, and thereby provide no additional clinical utility. In prostate cancer, Gleason sum remains the single most prognostic feature of localized cancer. Although many molecular assays have also been found to be associated with prognosis in prostate cancer, none are in broad use because most offer little to no independent prognostic information over Gleason sum and thus lack clinical utility.23

**Regulation of Molecular Diagnostic Tests**

The regulation of molecular testing falls under the jurisdiction of 2 federal agencies—the FDA and the Centers for Medicare and Medicaid Services (CMS). The laws and policies concerning the development and implementation of molecular tests continue to evolve. This section briefly reviews the key elements of the regulatory environment. NCCN also assembled a Molecular Testing Working Group in 2011, which generated a White Paper detailing the regulatory environment and highlighting additional challenges to the development and implementation of biomarkers.24

**FDA**

Medical tests are regulated by the Center for Devices and Radiological Health, a different part of the FDA from that which regulates drugs (Center for Drug Evaluation and Research) or biologics (Center for Biologics Evaluation and Research), leading to regulatory differences that impact FDA evaluation and approval. Regulatory evaluation of medical devices began with the 1976 Medical Device Amendment to the 1938 Federal Food, Drug, and Cosmetic Act, enacted in response to deaths and injuries attributable to some medical device products.25 Devices are divided into 3 classes based primarily on the risks associated with their intended use. For in vitro di-
agnostic tests, the risks are framed in terms of the consequences for patients when a test renders a wrong result. Class I tests are low risk, usually of a simple design, and typically are exempted from review by the FDA before proceeding to market. Class II tests pose moderate risk, are evaluated by the FDA through review of a 510(k) premarket notification, and are cleared for marketing once they are found to be substantially equivalent to a legally marketed device that was previously cleared by the FDA. Class III tests are those associated with the highest clinical risk. Each Class III test is reviewed through application for premarket approval (PMA). A demonstration of safety and effectiveness, which (for in vitro diagnostic devices) rests largely on a demonstration of analytic and clinical validity, is needed to gain FDA approval. The FDA approval/clearance process provides reasonable assurance of safety and effectiveness (i.e., that the benefit from using the test will outweigh the risk), and that the test will provide clinically significant results. The device regulations do not require a complete evaluation of clinical utility; hence, aspects such as cost-effectiveness or comparative effectiveness are not part of premarket review. A Class III device or test that has been approved for marketing through the PMA process is called FDA-approved, whereas a Class II device that has undergone premarket notification review through 510(k) is referred to as FDA-cleared.

The rapid translation of new scientific knowledge to medical practice has led to development and use of diagnostic tests that are often called home brew tests or laboratory developed tests (LDTs). Although the FDA retains jurisdiction over this category of test, until recently it practiced a policy of enforcement discretion for LDTs, choosing instead to regulate the critical reagents that are used to build these tests through the Analyte Specific Reagent Rule (21 CFR 864.4020, 809.10, and 809.30). Enforcement discretion has allowed more tests to be developed and performed, but has also come under recent scrutiny because of the lesser requirements for validation and reporting. The FDA recently released a preliminary guidance on Companion Diagnostics, or molecular tests which directly impact the use of a pharmaceutical or biologic drug, and the agency is expected to take a more active role in molecular test governance as more complex tests begin to enter clinical practice and are used to direct patient care. Importantly, for trials using Class III molecular assays to guide therapy, investigators are now often required to file an investigational device exemption with the FDA before initiating the trial. This relatively recent additional regulatory requirement ensures that acceptable standards of analytic validity are in practice but also places increased regulatory burden on investigators.

**CMS Clinical Laboratory Improvement Amendment**

Clinical laboratories are governed through the Clinical Laboratory Improvement Amendment (CLIA) of 1988 administered by CMS. **CLIA certification** is a descriptor that is frequently associated with analytically valid assays performed in a clinical laboratory and/or a laboratory that has received CLIA certification. Importantly, individual clinical laboratory assays are not CLIA-certified. Rather, CLIA certification requires that a laboratory adopt specific practices and perform prescribed measures of analytic validation while performing specific assays. CLIA certification is required for laboratories performing low-complexity waived tests and moderate- or high-complexity tests on clinical samples, whether they are FDA-approved, FDA-cleared, or laboratory developed. It has become clear that assays must be performed in CLIA-certified laboratories if the results of the assays are going to be used to guide patient management, regardless of FDA status. CLIA certification prescribes a certain level of analytic validation, but performance in a CLIA-certified laboratory does not address the clinical validity or clinical utility of an assay. However, FDA clearance or approval of a marker has been mostly based on analytic and clinical validity, but not necessarily clinical utility. Therefore, FDA approval does not necessarily mean that a marker should be used, and lack of FDA approval does not mean it should not be used, because an assay with clinical utility can be performed in a CLIA-certified laboratory without FDA approval. Key components of the regulation and oversight of CLIA-accredited laboratories are inspections and proficiency tests. In most jurisdictions and hospitals, these are provided by the College of American Pathologists (CAP). CAP inspection guidelines essentially set the standards for overall biomarker laboratory operation, including assay validation, quality control, and quality assurance activities. The guidelines mandate regular proficiency testing for every
clinical assay, and CAP runs an extensive program providing proficiency testing samples for all commonly used clinical assays.

The FDA recently withdrew its enforcement discretion for certain LDTs that are marketed “direct to consumers,” and previously published a draft guidance proposing increased regulation of tests described as In Vitro Diagnostic Multivariate Index Assays. The recently released draft guidance on Companion Diagnostics indicates the FDA’s intent to ensure the availability of safe and effective tests when results from those tests are essential for the safe and effective use of a corresponding therapeutic product (drug, biologic, or device).

### Systems for Level of Evidence

Ideally, clinical utility of a marker should be determined in a prospective clinical trial, as is required for new drugs. In fact, trial designs to determine clinical utility of markers have been described and several are underway. One way to streamline the evaluation of both established and upcoming molecular tests, using the guidelines discussed earlier to evaluate the quality of studies, is to assign a well-defined level of evidence to each test, using standards of analytic and clinical validity and clinical utility. Levels-of-evidence standards for assessing tumor markers have been published, and provide a common comparison that can be used across all types of molecular testing. Two in particular, the Tumor Marker Utility Grading System (TMUGS; Table 1) and the levels of evidence standards for using archived tissue (Table 2), have been proposed to provide a common language to improve communication regarding assay development and validation.

The authors of TMUGS proposed an updated annotation for levels of evidence that combines both Tables 1 and 2, while acknowledging the importance of rigorous analysis of archived specimens from definitive clinical studies (Table 3). The authors appropriately point out that biomarker development can challenge the semantic distinction between prospective and retrospective studies, and the critical element to biomarker validation is the robust statistical design of the study and definition of the biomarker parameters before application to the samples being used for validation. Although prospective collection of these samples in the setting of a clinical trial with objectives focused on the validation of a biomarker will remain the best means to minimize bias and obtain the highest level of evidence in support of a biomarker, robust analysis of archived specimens can also provide strong validation data. An example of the latter is the association of KRAS mutations and resistance to antibodies targeting epidermal growth factor receptors (EGFRs).

Thus, this revised levels of evidence system combines the original TMUGS system in Table 1 with the clinical trial design classification from Table 2 to create a system that can be used to classify tumor markers and also suggest what supportive validation studies are required for each level.

Although the evolution of systems with which to classify levels of evidence for biomarkers is likely to continue, the task force members believe that these level-of-evidence systems provide an appropriate initial mechanism of assigning levels of evidence for new assays that is consistent with how NCCN approaches the level of evidence and consensus for cancer therapies (Table 4). In the following sections that focus on specific diseases, the task force used the revised system proposed by Simon et al. to annotate the level of evidence supporting each tumor marker, and has noted the NCCN category of evidence and consensus for the markers included within the NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines; see Tables 5–10).

### Assays Currently Used in Cancer Care

For each disease considered by the task force, including glioma, breast cancer, colon cancer, lung cancer, prostate cancer, and acute myelogenous...
leukemia, tabular data were assembled on markers that are currently included in the NCCN Guidelines (available at www.NCCN.org), and new markers currently under investigation. For all listed biomarkers, NCCN categories of evidence and consensus were given if available, and combined levels of evidence calculations (Tables 1–3) were made and are included with references to illustrate the hurdles that prevent some new tests from being widely included in practice guidelines. The hope is that a critical examination of the published evidence on each of these tests will show why some, which may have been widely presented at meetings or advertised in public venues, do not achieve thresholds needed for broad adoption into clinical practice.

**Molecular Testing in Glioma**

Table 5 summarizes current molecular biomarkers in glioma.

**1p/19q Deletion**

The current NCCN Guidelines for Central Nervous System Cancers specify a single molecular test to help determine prognosis and select chemotherapy

---

**Table 2 Use of Archived Tissues to Determine Clinical Validity of Tumor Markers**

<table>
<thead>
<tr>
<th>Category</th>
<th>A Prospective</th>
<th>B Prospective Using Archived Samples</th>
<th>C Prospective/ Observational</th>
<th>D Retrospective/ Observational</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical trial</td>
<td>PCT designed to address tumor marker</td>
<td>Prospective trial not designed to address tumor marker, but design accommodates tumor marker utility</td>
<td>Prospective observational registry, treatment and follow-up not dictated</td>
<td>No prospective aspect to study</td>
</tr>
<tr>
<td>Patients and patient data</td>
<td>Prospectively enrolled, treated, and followed in PRCT</td>
<td>Prospectively enrolled, treated, and followed up in clinical trial and, especially if a predictive utility is considered, a PRCT addressing the treatment of interest</td>
<td>Prospectively enrolled in registry, but treatment and follow-up standard of care</td>
<td>No prospective stipulation of treatment or follow-up; patient data collected through retrospective chart review</td>
</tr>
<tr>
<td>Specimen collection, processing, and archival</td>
<td>Specimens collected, processed, and assayed for specific marker in real time</td>
<td>Specimens collected, processed, and archived prospectively using generic SOPs; assayed after trial completion</td>
<td>Specimens collected, processed, and archived prospectively using generic SOPs Assayed after trial completion</td>
<td>Specimens collected, processed, and archived with no prospective SOPs</td>
</tr>
<tr>
<td>Statistical design and analysis</td>
<td>Study powered to address tumor marker question</td>
<td>Study powered to address therapeutic question and underpowered to address tumor marker question Focused analysis plan for marker question developed before performing assays</td>
<td>Study not prospectively powered at all; retrospective study design confounded by selection of specimens for study Focused analysis plan for marker question developed before performing assays</td>
<td>Study not prospectively powered at all; retrospective study design confounded by selection of specimens for study No focused analysis plan for marker question developed before performing assays</td>
</tr>
<tr>
<td>Validation</td>
<td>Result unlikely to be play of chance Although preferred, validation not required</td>
<td>Result more likely to be play of chance than A, but less likely than C Requires one or more validation studies</td>
<td>Result very likely to be play of chance Requires subsequent validation studies</td>
<td>Result very likely to be play of chance Requires subsequent validation studies</td>
</tr>
</tbody>
</table>

Abbreviations: PCT, prospective controlled trial; PRCT, prospective, randomized controlled trial; SOP, standard operating procedure.

Other molecular tests are rapidly gaining attention in central nervous system cancers and, with additional clinical data and validation, will most likely be considered for inclusion in upcoming treatment guidelines. One of these is mutation in the cytosolic isocitrate dehydrogenase gene (\textit{IDH1}) or the mitochondrial version of the same gene (\textit{IDH2}). \textit{IDH1} is involved in the metabolic conversion of isocitrate to alpha-ketoglutarate, which reduces NADP to NADPH. In glioma, mutations in this gene were discovered through large-scale sequencing of tumor DNA samples.\(^48\) \textit{IDH1} mutations were present in 12% of glioblastoma samples in this study. Later retrospective analyses of 271 clinical trial samples of low-grade glioma showed that mutated \textit{IDH1} may be both prognostic and predictive, because it is associated with longer survival times and better response to temozolomide.\(^49,50\) Another study of astrocytoma showed an association with improved survival but not with response to temozolomide.\(^51\) Among WHO grade II and III gliomas, 50% to 80% have mutated \textit{IDH1}, whereas 5% to 10% of WHO grade IV gliomas bear a mutated \textit{IDH1}. Roughly 90% of \textit{IDH1} mutated proteins can be detected using immunohistochemistry with a monoclonal antibody that detects p.R132H, the most common \textit{IDH1} mutation.\(^52,53\)

A recommendation for optimal testing for this mutation might be to first perform immunohistochemical testing with this detecting antibody, then to follow up with DNA sequencing only when the results from immunohistochemistry are negative or equivocal. Further clarification of the correlation of chemotherapy response and \textit{IDH1} mutation will be needed for this test to receive broad clinical use. At this point, lack of \textit{IDH1} mutation could not be considered strong enough evidence to alter for certain types of central nervous system tumors. Testing for the 1p/19q codeletion, or unbalanced translocation, is recommended for cases of suspected oligodendroglioma. The test is also recommended to distinguish anaplastic oligodendroglioma from anaplastic astrocytomas and glioblastomas (to view the most recent version of these guidelines, visit the NCCN Web site at www.NCCN.org).\(^35\)

Oligodendroglioma displays a loss of 1p/19q and is more sensitive to chemotherapy and radiotherapy than are astrocytic tumors. Presence of the translocation is a favorable prognostic factor, and has an NCCN category 1 designation, indicating that high-level evidence, either from randomized controlled clinical trials or meta-analyses, supports using this test.\(^36-42\) Detection methods, not specified within the guideline, include polymerase chain reaction (PCR)–based testing for loss of heterozygosity, FISH, array comparative genomic hybridization, and multiplex ligation-dependant probe amplification (MPLA).\(^42-46\) A practical method of FISH testing for 1p/19q deletion was recently published\(^47\) and the authors note that no other consensus guidelines or protocols exist for 1p/19q testing using other technologies. This test is considered to have level 1A evidence (Table 3).

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|}
\hline
Level of Evidence & Category From Table 2 & Validation Studies Available \\
\hline
I & A & None required \\
II & B & One or more with consistent results \\
II & C & None or inconsistent results \\
III & C & 2 or more with consistent results \\
IV–V & D & None or 1 with consistent results or inconsistent results \\
\hline
\end{tabular}
\caption{Revised Determination of Levels of Evidence Using Elements of Tumor Marker Studies*}
\end{table}

*Levels of evidence revised from those originally proposed in Tables 1 and 2.\(^31\)

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|}
\hline
Category & From Table 2 & Validation Studies Available \\
\hline
I A & None required \\
I B & One or more with consistent results \\
II B & None or inconsistent results \\
II C & 2 or more with consistent results \\
III C & None or 1 with consistent results or inconsistent results \\
IV–V D & NA\(^1\) \\
\hline
\end{tabular}
\caption{NCCN Categories of Evidence and Consensus}
\end{table}

\begin{flushleft}
\textbf{Table 3 Revised Determination of Levels of Evidence Using Elements of Tumor Marker Studies*}
\end{flushleft}

\begin{flushleft}
\textbf{Table 4 NCCN Categories of Evidence and Consensus}
\end{flushleft}

\begin{flushleft}
\textbf{IDH1/IDH2 Mutation}
\end{flushleft}

\begin{flushleft}
Other molecular tests are rapidly gaining attention in central nervous system cancers and, with additional clinical data and validation, will most likely be considered for inclusion in upcoming treatment guidelines. One of these is mutation in the cytosolic isocitrate dehydrogenase gene (\textit{IDH1}) or the mitochondrial version of the same gene (\textit{IDH2}). \textit{IDH1} is involved in the metabolic conversion of isocitrate to alpha-ketoglutarate, which reduces NADP to NADPH. In glioma, mutations in this gene were discovered through large-scale sequencing of tumor DNA samples.\(^48\) \textit{IDH1} mutations were present in 12% of glioblastoma samples in this study. Later retrospective analyses of 271 clinical trial samples of low-grade glioma showed that mutated \textit{IDH1} may be both prognostic and predictive, because it is associated with longer survival times and better response to temozolomide.\(^49,50\) Another study of astrocytoma showed an association with improved survival but not with response to temozolomide.\(^51\) Among WHO grade II and III gliomas, 50% to 80% have mutated \textit{IDH1}, whereas 5% to 10% of WHO grade IV gliomas bear a mutated \textit{IDH1}. Roughly 90% of \textit{IDH1} mutated proteins can be detected using immunohistochemistry with a monoclonal antibody that detects p.R132H, the most common \textit{IDH1} mutation.\(^52,53\) A recommendation for optimal testing for this mutation might be to first perform immunohistochemical testing with this detecting antibody, then to follow up with DNA sequencing only when the results from immunohistochemistry are negative or equivocal. Further clarification of the correlation of chemotherapy response and \textit{IDH1} mutation will be needed for this test to receive broad clinical use. At this point, lack of \textit{IDH1} mutation could not be considered strong enough evidence to alter
therapy. However, one study showed that IDH wild-type grade III tumors (usually treated with radiotherapy alone) are associated with a prognosis similar to grade IV tumors (treated with concurrent chemotherapy and radiotherapy). If validated with further studies, IDH testing could therefore have an impact on patient management. Because the presence of an IDH mutation is considered tumor-specific, it has a role as a diagnostic marker when morphologic features are inconclusive and a nonneoplastic/reactive condition is possible. Overall, this marker is currently considered primarily prognostic/diagnostic, with the possibility that more data will indicate whether mutation status can be considered in recommending treatment. Studies showing the association have come from well-designed retrospective analyses, and evidence for the test is considered to be level IIB.

**BRAF Fusion Protein**

Alterations in the signaling-associated kinase BRAF have been noted in many tumors. These often occur as point mutants which lead to a constitutively active enzyme, BRAF c.1799T>A (p.V600E). Most pilocytic astrocytomas have been shown to express a fusion protein of KIAA1549:BRAF, which is constitutively active. Because pilocytic astrocytoma can be cured through surgery alone, testing for the presence of the fusion protein might be useful to aid in the diagnosis of pilocytic astrocytoma when histologic features are inconclusive and difficult to distinguish from a diffuse astrocytoma. Therefore the combination of IDH mutation testing (common in diffuse astrocytoma) with BRAF testing represents a rational diagnostic panel for cases with indeterminate histologic features. Testing methodologies used in research studies include long-distance inverse PCR to map breakpoints on the genomic level, 5′ rapid amplification of cDNA ends (RACE) to detect gene fusion, and FISH using probes that map to the KIAA1549 sequence and to the BRAF sequence. Implementation of the FISH assay will probably be the most straightforward for clinical testing, as this can be performed on FFPE samples, but lack of commercially available probe sets and assay standardization.

### Table 5  Current Molecular Biomarkers in Glioma

<table>
<thead>
<tr>
<th>Biomarker</th>
<th>Molecular Compartment</th>
<th>Purpose</th>
<th>Analytic Validity Demonstrated</th>
<th>Level of Evidence</th>
<th>NCCN Category of Evidence</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Markers With Accepted Clinical Utility</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1p/19q codeletion (unbalanced translocation)</td>
<td>Tumor DNA</td>
<td>Diagnostic (oligodendroglioma)</td>
<td>FISH, aCGH, LOH, MPLA</td>
<td>IA</td>
<td>1</td>
<td>Smith et al.46</td>
</tr>
<tr>
<td>IDH mutation (IDH1) c. 395 G&gt;A p.R132H (IDH2)</td>
<td>Tumor DNA, tumor protein</td>
<td>Positive is favorably prognostic; also a diagnostic marker</td>
<td>IHC, DNA sequencing</td>
<td>IIB</td>
<td></td>
<td>Houillier et al.49 Dubbink et al.31</td>
</tr>
<tr>
<td>MGMT methylation</td>
<td>Tumor DNA</td>
<td>Prognostic, predictive (benefit for chemotherapy), pharmacodynamic (pseudorecurrence)</td>
<td>MS-PCR, MS-pyrosequencing, MS-MPLA</td>
<td>IIB</td>
<td></td>
<td>Hegi et al.51 Gilbert et al.215</td>
</tr>
<tr>
<td><strong>Markers With Emerging Evidence</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BRAF fusion (pilocytic astrocytoma)</td>
<td>Tumor DNA</td>
<td>Diagnostic (pilocytic astrocytoma)</td>
<td>LDI-PCR, 5′ RACE, FISH</td>
<td>IIB</td>
<td></td>
<td>Jeuken and Wesseling.216 Jones et al.59</td>
</tr>
<tr>
<td>CIMP (CpG island methylator phenotype)</td>
<td>Tumor DNA</td>
<td>Positive is favorably prognostic</td>
<td>Gene expression microarray, pyrosequencing</td>
<td>IIB</td>
<td></td>
<td>Noushmehr et al.55 Gilbert et al.215</td>
</tr>
</tbody>
</table>

Abbreviations: aCGH, array comparative genomic hybridization; FISH, fluorescence in situ hybridization; IHC, immunohistochemistry; LDI-PCR, long-distance inverse polymerase chain reaction; LOH, loss of heterozygosity; MPLA, multiplex ligation-dependant probe amplification; MS, methylation-specific; MS-MPLA, methylation-specific multiplex ligation-dependant probe amplification; MS-PCR, methylation-specific polymerase chain reaction; RACE, rapid amplification of cDNA ends.
tion to date still must be overcome for this test to be broadly implemented. The association has been shown through well-designed retrospective analyses, and evidence for the test is considered to be level IIB.

O-6-Methylguanine-DNA-Methyltransferase Promoter Methylation
Activity of DNA repair enzyme O-6-methylguanine-DNA methyltransferase (MGMT) is associated with resistance to chemotherapeutic alkylating agents. Silencing of the MGMT promoter through methylation has been shown to be associated with better outcomes (overall survival, progression-free survival) in a variety of glioma subtypes and to partially predict response to alkylating agents. MGMT promoter methylation has also been shown to be positively associated with therapeutic pseudo-regression. One barrier to use of MGMT as a prognostic marker, or as a marker for pseudo- versus true regression, is that the methodology for assessing MGMT is not yet standardized. Measurement of MGMT protein levels using immunohistochemical testing does not correlate with extent of MGMT promoter methylation as measured using methylation-specific PCR. Promoter methylation itself can be measured using a variety of technologies, including methylation-specific PCR, methylation-specific pyrosequencing, and methylation-specific MPLA. Each of these assays has benefits and drawbacks in terms of reproducibility, standardization of cutoffs, and ease of use, and which will emerge as the best for use in a clinical setting remains to be seen. The association has been shown through well-designed retrospective analyses, and evidence for the test is considered to be level IIB.

Glioma-CpG Island Methylator Phenotype
A more-extensive methylation analysis of glioblastoma tumors showed a glioma-CpG island methylator phenotype (G-CIMP) similar to that observed in colorectal cancer. The G-CIMP phenotype is colocalized with the IDH1 mutation, and is found predominately in lower-grade tumors from younger patients with longer survival times. Promoter methylation is observed in a subset of loci, defining a set of glioma tumors. Methods involved assessing promoter methylation and gene expression for a training set of tumors, and assembling a list of genes that were both transcriptionally silenced and methylated. These results, though intriguing, will require further analytic validation, assay development, and especially demonstration that the G-CIMP phenotype is associated with greater or lesser sensitivity to treatment regimens. The association has been shown through well-designed retrospective analyses, and evidence for the test is considered to be level IIB (Tables 1–3).

Overall, use of central nervous system tumor biomarkers is in its infancy, with several new and exciting candidates that have the potential to change how tumors are evaluated and may help determine treatment. The challenge in this area will be to ensure that clinical recommendations encompass validated assays, including specific recommendations of methodology and scoring, and include comparisons of technologies used to assess these biomarkers.

Molecular Testing in Breast Cancer
Table 6 summarizes current molecular biomarkers in breast cancer.

*ER*-α(ERα1)/PgR(PR)
The measurement of estrogen receptor expression is prognostic and predictive for response to *ER*-α-modulating agents. The prognostic and treatment predictive use of progesterone receptor expression is less clear. Testing for *ER*-α (ERα1) in ductal carcinoma in situ and *ER*-α(ERα1)/PgR(PR) expression in breast tumors is a necessary component of the initial workup for these diseases. Guidelines for immunohistochemical testing in breast cancer have been published by ASCO, CAP, and NCCN, with detailed information available on all aspects of testing, including tissue handling and assay validation. The NCCN category of evidence and consensus for use of this test is 2A. Using the TMUGS and Simon et al. criteria for assessing marker validation gives this test a rating of IB as a predictive test (Table 3).

HER2(ERBB2)
Expression of HER2(ERBB2) in breast tumors is also both a prognostic and predictive marker. HER2(ERBB2)-positive tumors respond to trastuzumab, a monoclonal antibody directed against the extracellular domain of the HER2(ERBB2) protein, and lapatinib, a tyrosine kinase inhibitor, versus HER2(ERBB2)-negative tumors which do not seem to respond. Clinical trial results have shown that adjuvant trastuzumab reduces recurrence and mortality by approximately 50% and 30%, respec-
In HER2(ERBB2)-positive patients,10–12 and inclusion of trastuzumab in the therapy regimen for HER2(ERBB2)-positive patients is given a category 1 recommendation in the NCCN Guidelines for Breast Cancer,72 with a combined level of evidence score of 1A. Accurate testing for HER2(ERBB2) is thus critically important in developing treatment plans for breast cancer. NCCN and ASCO/CAP have both made extensive and detailed guidelines on HER2(ERBB2) testing for both FISH and immunohistochemistry,77,78 and the NCCN Guidelines for Breast Cancer contain a section devoted to this issue72 (available online, at www.NCCN.org [BINVAJ]), which details specific thresholds and confirmation of positive and negative results.

### Oncotype Dx

The Oncotype Dx (Genomic Health, Inc., Redwood City, CA) test is a 21-gene reverse transcriptase PCR (RT-PCR) assay that is performed at a single site as a laboratory-developed test, using FFPE breast tumor samples as a source of RNA. Gene expression levels are translated into a recurrence score.79–81 The current NCCN Guidelines recommend the 21-gene recurrence score be considered for determining prognosis in node-negative, ER-α(ESR1)-positive breast tumors. Patients with a low recurrence score have such a favorable prognosis that even if chemotherapy were beneficial, so few patients would benefit that the risks outweigh the benefits. The NCCN category of evidence and consensus for using Oncotype DX to determine whether a patient receives chemotherapy is 2B.

---

### Table 6 Current Molecular Biomarkers in Breast Cancer

<table>
<thead>
<tr>
<th>Biomarker</th>
<th>Molecular Compartment</th>
<th>Purpose</th>
<th>Analytic Validity Demonstrated</th>
<th>Level of Evidence</th>
<th>NCCN Category of Evidence</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Markers With Accepted Clinical Utility</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ER-α/PgR (ESR1/PR)</td>
<td>Tumor protein</td>
<td>Diagnostic (weak) predictive</td>
<td>IHC</td>
<td>Predictive: IB</td>
<td>2A</td>
<td>Early Breast Cancer Trialists Collaborative Group68 Hammond et al.74 Hammond et al.75 Allred et al.73</td>
</tr>
<tr>
<td>HER2(ERBB2)</td>
<td>Tumor protein</td>
<td>Diagnostic (classification) prognostic (favorable) predictive for anti-HER2(ERBB2) therapy</td>
<td>FISH, IHC</td>
<td>Predictive: IA</td>
<td>2A</td>
<td>Wolff et al.76 Carlson et al.77 Joensuu et al.70 Piccart-Gebhart et al.11 Romond et al.72</td>
</tr>
<tr>
<td>Oncotype Dx</td>
<td>Tumor RNA</td>
<td>Prognostic predictive</td>
<td>21-gene RT-PCR expression assay of FFPE samples with recurrence score as readout</td>
<td>Prognostic in ER+/node-negative patients: IB Predictive: IIA</td>
<td>Selection of treatment options based on Oncotype Dx score is category 2B, but use of the test in a defined subpopulation of patients is 2A</td>
<td>Paik et al.80 Paik et al.81 Albain et al.79</td>
</tr>
</tbody>
</table>

**Markers With Emerging Evidence**

<table>
<thead>
<tr>
<th>Biomarker</th>
<th>Molecular Compartment</th>
<th>Purpose</th>
<th>Analytic Validity Demonstrated</th>
<th>Level of Evidence</th>
<th>NCCN Category of Evidence</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>MammaPrint</td>
<td>Tumor RNA</td>
<td>Not clear</td>
<td>Multigene microarray expression assay using frozen tissue</td>
<td>IIB</td>
<td></td>
<td>Straver et al.217 Knauer et al.218</td>
</tr>
<tr>
<td>CTCs</td>
<td>Circulating tumor cells</td>
<td>Monitor patients with metastatic disease</td>
<td>Cell surface staining and magnetic separation</td>
<td>IB</td>
<td></td>
<td>Budd et al.219 Cristofanilli et al.90,220 De Giorgi et al.221</td>
</tr>
</tbody>
</table>

Abbreviations: CTC, circulating tumor cell; ER, estrogen receptor; FFPE, formalin-fixed paraffin-embedded; FISH, fluorescence in situ hybridization; IHC, immunohistochemistry; RT-PCR, reverse transcriptase polymerase chain reaction.
This recommendation is echoed by ASCO in its 2007 recommendations for use of tumor markers in breast cancer. In its review of tumor gene expression profiling in breast cancer, the Evaluation of Genomic Applications in Practice and Prevention (EGAPP) working group found insufficient evidence to recommend Oncotype Dx but noted that encouraging indirect evidence exists for clinical utility of the test.

Provocative but still hypothesis-generating studies have suggested that in addition to being prognostic in women with ER-α (ESR1)-positive, node-negative tumors, the 21-gene recurrence score may also be predictive of the relative benefits of chemotherapy. These studies have suggested that tumors with low recurrence scores are relatively resistant to chemotherapy, whereas those with higher recurrence scores are increasingly sensitive, independent of the prognostic effect of the assay.

The recently completed TAILORx trial will determine the benefit of chemotherapy in node-negative women with intermediate risk scores, and will further determine the prognostic role of the assay in patients with a low recurrence score. The ongoing S01007/RxPONDER trial will prospectively determine the benefits of adjuvant chemotherapy in node-positive, ER-α (ESR1)-positive patients. As trial results become available, and if they continue to show a demarcation in value for chemotherapy, this assay is expected to become more widely recommended. Currently, the level of evidence score for Oncotype Dx as a prognostic test is IB, whereas as a predictive test, it is IIA.

**MammaPrint**

The MammaPrint 70-gene microarray assay (Agendia, Irvine, CA) is the only multigene breast cancer assay that has received FDA clearance. However, 3 major groups that have considered it have not recommended its use. It is useful to remember that the FDA does not make considerations of clinical utility, only of analytic validity and clinical validity in its determinations of clearance. One barrier to the use of MammaPrint is the requirement of fresh-frozen tissue samples, rather than the more universally available FFPE tumor samples. In addition, currently completed studies do not assess the ability of MammaPrint to predict benefit from treatment. An ongoing trial (MINDACT) will compare the effectiveness of MammaPrint results to typical clinical evaluation in the prediction of survival at 15 years. Until trial results are available, MammaPrint has a combined level of evidence of IIB as a prognostic test.

**Circulating Tumor Cells**

The CellSearch circulating tumor cell (CTC) detection system (Veridex, LLC, Raritan, NJ) has been cleared for use in metastatic breast cancer by the FDA. However, this test is not recommended by the ASCO breast cancer tumor marker panel and is not included in the NCCN Guidelines for Breast Cancer. The test uses an immunomagnetic system to isolate cells with the features of epithelial cells from the peripheral blood and provides an absolute count of CTCs per 7.5-mL blood sample. Presence of greater than 5 CTCs per 7.5-mL sample has been shown to be associated with shortened progression-free (2.7 vs. 7 months) and overall survivals (10.1 vs. > 18 months). A randomized clinical trial, SWOG S0500, is designed to determine whether elevated CTC levels after initiation of treatment indicate that a change in chemotherapy regimen will be beneficial in the metastatic setting. Results of this trial will add to the clinical utility of this assay, and may allow it to be further considered by guidelines panels. Until those results are available, clinical validity for the determination of CTCs in breast cancer has a combined level of evidence score of IB.

**Molecular Testing in Colon Cancer**

Table 7 summarizes current molecular biomarkers in colon cancer.

**KRAS Mutations**

The RAS/RAF/MAPK and downstream pathways are activated in response to stimulation of the EGFR, and mutations to proteins in this pathway have been investigated for their role in modulating response to EGFR-targeting agents. Approximately 40% of colon cancers are positive for mutations in KRAS, and colon cancers with KRAS mutations in codons 12 and 13 have been shown to be unresponsive to EGFR-targeting monoclonal antibody therapies, such as cetuximab or panitumumab. The NCCN Guidelines for Colon Cancer recommend, with a category of 2A, the testing of all metastatic disease for the presence of mutations in KRAS (to view the most recent version of these guidelines, visit the NCCN Web site at www.NCCN.org).
ASCO provisional clinical opinion\textsuperscript{98} also recommends KRAS mutation testing in metastatic colon cancers. No FDA-approved KRAS mutation test exists, and both publications mention the necessity of performing the test in a CLIA-certified facility. The ASCO Provisional Clinical Opinion notes that the test can be performed using either real-time PCR with mutation-specific primers, or sequencing. The NCCN Guidelines do not recommend any particular methodology. Testing for KRAS mutations can also be given a combined level of evidence of IB, owing to multiple randomized controlled trials of panitumumab or cetuximab in colon cancer, with evaluation of KRAS mutation status in archived samples.\textsuperscript{33,34,96} Recent reports have shown that tumors harboring KRAS mutations at the c.38G>A (p.G13D) position may still be responsive to cetuximab, unlike other activating codon 12 and 13 mutations.\textsuperscript{99,103} The NCCN Guidelines note that the use of EGFR-targeting agents in KRAS codon 13 mutant tumors is investigational, and is not recommended for routine clinical practice at this time.\textsuperscript{97}

**BRAF c.1799T>A (p.V600E) Mutation**

The BRAF c.1799T>A (p.V600E) mutation renders the kinase domain of the protein constitutively active, and is found in 10% to 20% of colon tumors.\textsuperscript{96–98,102} RAF is downstream of EGFR, and a constitutively activated RAF enzyme might therefore circumvent EGFR-directed therapy. Retrospective studies have shown that mutated BRAF is associated with decreased progression-free and overall survivals in patients undergoing anti-EGFR–containing therapy,\textsuperscript{103–105} although interpretation of this association is confounded by its strong negative prognostic implications. Randomized data suggest, however, that patients with a BRAF mutation may benefit from EGFR-directed therapy in the first-line treatment setting, despite the unfavorable prognosis conferred by this mutation.\textsuperscript{95} The NCCN Guidelines for Colon Cancer (available at www.NCCN.org) recommend that BRAF testing be considered in the metastatic setting if the KRAS gene is found to be unmutated. This is given a category of evidence of 2A.\textsuperscript{97} Combined level of evidence for the use of BRAF testing for prognostic information is IB, owing to multiple retrospective analyses of randomized clinical trial samples.\textsuperscript{58,95,103–107}

Mutations of KRAS and BRAF almost never coexist in the same tumor.\textsuperscript{98} Testing for BRAF is recommended to be performed in a CLIA-certified laboratory, and is usually accomplished using Sanger sequencing, which can be performed with FFPE samples containing at least 40% neoplastic cellularity (~20% mutant alleles); detection limit can be 5% to 10% mutant alleles or lower in pyrosequencing and real-time PCR with or without melting curve analysis.\textsuperscript{108–110} Testing for BRAF mutation in colon cancer is primarily through laboratory-developed tests, because, other than the test recently approved for melanoma, no FDA-cleared BRAF mutation tests are currently available.

**Microsatellite Instability**

Defects in mismatch repair (MMR) can be detected using either immunohistochemistry for the absence of MLH1, MSH2, MSH6, and/or PMS2 protein in tumor tissues, or through a comparative PCR analysis of tumor and normal patient DNA for microsatellite instability (MSI). Hereditary syndromes with a predisposition to the development of colon cancer, such as Lynch syndrome, have been shown to contain germline mutations in one of the MMR genes, whereas most sporadic MSI-high colon tumors show MLH1 promoter hypermethylation and epigenetic silencing.\textsuperscript{102,109,111,112} In general, MSI-high colon cancers are associated with good prognosis, and patients with MSI-high stage II colon cancers do not seem to benefit from adjuvant single-agent 5-FU chemotherapy.\textsuperscript{113–115}

The NCCN Guidelines currently recommend testing for MMR proteins in all patients younger than 50 years and/or who meet the revised Bethesda guidelines or Amsterdam criteria because of the increased possibility of Lynch syndrome in this cohort, and with a level 2A category of evidence. MMR/MSI testing is also recommended for patients with stage II tumors if single-agent fluoropyrimidine adjuvant therapy is considered, given evidence for lack of benefit and possibly worse outcomes in this population (NCCN Guideline, category 2A).\textsuperscript{97} Use of MSI/MMR testing as a screen for Lynch syndrome has a combined level of evidence score of IB,\textsuperscript{116–119} as does use of this test for prognostic purposes,\textsuperscript{106,120,121} owing to the results of multiple studies, including both randomized clinical trials and meta-analyses. The ASCO recommendation for use of tumor markers in colorectal cancer does not endorse routine testing for MSI or MMR in all patients,\textsuperscript{122} but this recommendation may be reconsidered when the position statement is updated. No FDA-cleared tests are available for MSI testing.
Immunohistochemistry testing for MMR proteins has been shown to be a reliable method for detecting defects in the DNA MMR system, and is generally concordant with MSI testing with PCR.\textsuperscript{117,123} Germline mutation (plus a second hit) or silencing through promoter methylation both lead to an absence of functional protein, detectable through immunohistochemistry in most instances. The presence of a BRAF mutation in addition to the finding of MMR deficiency and/or MSI-high indicates a sporadic acquisition and may be used to distinguish from the hereditary Lynch syndrome.\textsuperscript{124,125} Reflexive testing of all newly diagnosed colorectal cancers for the presence of MMR deficiency and/or MSI, along with BRAF mutational status if MMR deficiency and/or MSI-high status are detected, may be recommended in the future as part of standard pathologic review pending guideline updates from CAP.

**CEACAM5 (CEA)**

Testing for carcinoembryonic antigen–related cell adhesion molecule 5, or CEACAM5 (so-called CEA), which is released from tumor cells into patient serum, is recommended by the NCCN Guidelines for Colon Cancer and ASCO.\textsuperscript{97,122} Testing is recommended on di-

### Table 7 Current Molecular Biomarkers in Colon Cancer (Cont. on facing page)

<table>
<thead>
<tr>
<th>Biomarker</th>
<th>Molecular Compartment</th>
<th>Purpose</th>
<th>Analytic Validity Demonstrated</th>
<th>Levels of Evidence</th>
<th>NCCN Category of Evidence</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>K\textsuperscript{RAS} mutations (\text{except c.38G}\to\text{A (p.G13D)})^\text{*}</td>
<td>Tumor DNA</td>
<td>Predictive (negative for anti-EGFR therapy); negatively prognostic in several first-line randomized studies</td>
<td>Multiple methods: PCR, multiplex assays, direct sequencing</td>
<td>Predictive: IB</td>
<td>Prognostic: IIB</td>
<td>2A</td>
</tr>
<tr>
<td>MSI and/or MMR protein loss</td>
<td>Tumor DNA for MSI testing with PCR; tumor IHC for MMR proteins</td>
<td>Screening (Lynch syndrome) Prognostic (recurrence, overall survival) Predictive (lack of benefit, possibly worse outcome with adjuvant single-agent fluoropyrimidine therapy)</td>
<td>PCR, IHC</td>
<td>Screening: IB</td>
<td>Prognostic: IB Predictive: IIB</td>
<td>2A</td>
</tr>
<tr>
<td>CEACAM5 (CEA)</td>
<td>Patient serum</td>
<td>Surveillance</td>
<td>Immunoassay</td>
<td>IIC</td>
<td>2A</td>
<td>Wanebo et al.\textsuperscript{126} Wang et al.\textsuperscript{127} Locker et al.\textsuperscript{122}</td>
</tr>
<tr>
<td>\textit{BRAF} c.1799T\to\text{A (p.V600E)} mutation</td>
<td>Tumor DNA</td>
<td>Prognostic (strong negative prognostic marker) Predictive (negative for anti-EGFR therapy)</td>
<td>Multiple methods: PCR, multiplex assays, direct sequencing</td>
<td>Prognostic: IB Predictive: IIIC</td>
<td>2A</td>
<td>Bokemeyer et al.\textsuperscript{95} De Roock et al.\textsuperscript{103} Ogino et al.\textsuperscript{58} Roth et al.\textsuperscript{106} Samowitz et al.\textsuperscript{107} Tol et al.\textsuperscript{105}</td>
</tr>
</tbody>
</table>

Abbreviations: CTC, circulating tumor cell; EGFR, epidermal growth factor receptor; FFPE, formalin-fixed paraffin-embedded; H&E, hematoxylin and eosin; IHC, immunohistochemistry; MMR, mismatch repair; MSI, microsatellite instability; PCR, polymerase chain reaction; RT-PCR, reverse transcriptase polymerase chain reaction.

\textsuperscript{*}All testing for \textit{K\textsuperscript{RAS}} mutation is presumed to have the same analytic validity, and will probably be performed simultaneously. However, presence of a c.38G\to\text{A (p.G13D)} mutation does not yet meet the standard of clinical utility needed to drive a treatment decision.
agnosis of invasive colon cancer, which can be used as a baseline for monitoring. CEACAM5 (CEA) is then tested regularly, and increasing levels are one indication of recurrent disease or metastasis. FDA-approved immunoassays are available for CEACAM5 (CEA) monitoring. Use of this test is a category 2A recommendation in the NCCN Guidelines, and the combined level of evidence score is IIC, because most studies on this marker have been observational.127,128,129

Neither guideline recommends CEACAM5 (CEA) as a screening or diagnostic tool.

CTCs
Detection of CTCs in patients with advanced colorectal cancer is a negative prognostic factor, correlating with reduced progression-free and overall survivals.128-130 The clinical utility of this test is unclear. CTC testing is not included in the NCCN Guidelines for Colon Cancer (to view the most recent version of these guidelines, visit the NCCN Web site at www.NCCN.org). As a prognostic tool, evidence for the clinical validity of CTC testing is level IB.

ColoPrint
The ColoPrint assay uses oligonucleotide microarray technology to assign a recurrence score based on expression of 18 genes, and currently requires fresh frozen tumor tissue as a source of RNA. Higher recurrence scores are associated with shorter time to progression and shorter overall survival in patients with stage I through III colon and rectal cancers, but...
are not predictive of chemotherapy benefit.\textsuperscript{131,132} The test can be considered to have developed clinical validity, with a IIC level of evidence, but does not have proven clinical utility at this time. It is not included in the NCCN Guidelines for Colon Cancer (available at www.NCCN.org).

**Oncotype Dx Colon Assay**

The Oncotype Dx Colon assay is a 12-gene RT-PCR expression assay that can be performed on FFPE tumor samples. This is also a prognostic test, with higher recurrence scores associated with shorter disease-free and overall survivals. The Oncotype Dx colon recurrence score has established clinical validity in its association with recurrence risk in patients with stage II colon cancer in 2 independent randomized validation studies, with a level of evidence of IB, but its clinical utility has not yet been established. It is not included in the NCCN Guidelines for Colon Cancer (available at www.NCCN.org).

**Other Biomarkers**

Other tests, including CIMP status,\textsuperscript{58,102,133,134} LINE-1 (global DNA) hypomethylation,\textsuperscript{135,136} and presence of infiltrating immune cells within colon tumors,\textsuperscript{137–142} are also under current investigation for their use in guiding colon cancer treatment. These tests will require further analytic and clinical validation, and demonstration of clinical utility before they should be considered for use in general practice.

**Molecular Testing in Non–Small Cell Lung Cancer**

Table 8 summarizes current molecular biomarkers in non–small cell lung cancer (NSCLC).

**EGFR Mutation**

The EGFR tyrosine kinase is expressed in most lung cancers, and in 2004, somatic kinase domain mutations that led to a constitutively activated EGFR were shown to be associated with high sensitivity to the small-molecule EGFR kinase inhibitors gefitinib and erlotinib.\textsuperscript{143–145} The most common mutations are exon 19 deletions (LREA deletion) and exon 21 point mutations (most commonly L858R),\textsuperscript{146} which together account for 85% to 90% of EGFR mutations in lung cancer. EGFR mutations are found more often in women, never-smokers, and East Asians.\textsuperscript{147} Because the presence of a mutation confers sensitivity to tyrosine kinase inhibition, and nonmutated EGFR tumors do not respond to EGFR inhibition, the NCCN Guidelines for NSCLC recommend testing recurrent or metastatic adenocarcinoma of the lung for EGFR mutation, with a category 1 recommendation and a combined level of evidence score of IA owing to multiple randomized clinical trials (to view the most recent version of these guidelines, visit the NCCN Web site at www.NCCN.org).\textsuperscript{148} In early 2011, ASCO published a preliminary clinical opinion,\textsuperscript{149} indicating that analysis of clinical trial data suggest a benefit for EGFR testing. A variety of methods are available for testing for EGFR mutations,\textsuperscript{150} and the current guidelines do not recommend a particular test, noting only that mutation testing is preferable to FISH assessment of EGFR copy number or immunohistochemistry to assess EGFR protein levels.\textsuperscript{151–153} A collaborative effort between CAP, the Association for Molecular Pathology (AMP), and the International Association for the Study of Lung Cancer (IASLC) is scheduled to produce detailed guidelines for EGFR and ALK testing in lung cancer, which should be a valuable resource for guideline developers and practitioners.

**KRAS Mutation**

Constitutively activating KRAS mutations are detected in approximately 20% of patients with NSCLC, mainly in those with adenocarcinoma and in smokers. Several studies analyzing both EGFR and KRAS mutations have shown that these mutations are present in mutually exclusive populations,\textsuperscript{154–156} and that the presence of a KRAS mutation is a negative prognostic factor in the absence of treatment. KRAS is a downstream component of the EGFR signaling pathway, and activating mutations in this protein might therefore be expected to be unaffected by EGFR inhibition. KRAS mutant tumors have been shown to be resistant to gefitinib or erlotinib therapy.\textsuperscript{157–159} However, the NCCN Guidelines for NSCLC (available at www.NCCN.org) currently do not recommend KRAS testing for any particular patient population because the impact on progression-free or overall survival is unclear,\textsuperscript{148} and the combined level of evidence score for KRAS testing is IIB. Notably, no one has performed a study randomizing patients with KRAS mutant tumors to erlotinib versus chemotherapy. However, a tumor that has a KRAS mutation is highly likely to be EGFR wild-type, and therefore the EGFR wild-type status can be used to inter-
pret use of EGFR tyrosine kinase inhibitors. Furthermore, KRAS mutations define a distinct subset of lung cancer that may be treated with different therapies in the future.

**ALK Gene Fusion**
A chromosomal rearrangement resulting in a fusion protein involving the ALK tyrosine kinase is found in 2% to 7% of patients with NSCLC. Most of the time, the 5′ fusion partner is EML4, but other 5′ partners, such as KIF5B and TFG, are rarely identified. Patients with ALK translocations are similar to those with EGFR mutations in that they are more often never-smokers with adenocarcinoma; however, similar to EGFR mutations, ALK fusions can be found in all types of lung adenocarcinoma, regardless of clinical features. For this group of patients, ALK inhibition represents a valuable therapeutic tool. The FDA recently approved crizotinib, a small molecule inhibitor of ALK, together with a companion diagnostic test for the translocation. FISH and immunohistochemistry have both been used to detect ALK fusion products, and some recommend confirming mutation status using both assays, but testing for ALK fusion using FISH was an enrollment criterion for the crizotinib trials, and this is the only FDA-approved ALK test available at this time.

Because ALK expression is lower in lung tumors than in anaplastic large cell lymphoma (which is also characterized by ALK gene fusions), a more sensitive immunohistochemistry test is needed for use in lung tumors. Testing for ALK fusions is included in the most recent NCCN Guidelines for NSCLC with a category 2A designation, and has a combined level of evidence of IIB.

**ERCC1 Expression**
ERCC1 is a component of the DNA excision repair pathway, and is required for the repair of DNA lesions introduced by ultraviolet light or formed by compounds such as cisplatin. High levels of ERCC1 expression in NSCLC, as measured with either RT-PCR or conventional immunohistochemistry, correlate with longer survival times in the absence of chemotherapy. However, tumors with low expression of ERCC1 do derive benefit from adjuvant cisplatin-based chemotherapy, whereas those with high levels of ERCC1 expression do not. The test, or any particular methodology, is not currently recommended in the NCCN Guidelines, and can be considered to have achieved a IIB level of evidence.

**Other Biomarkers**
Several additional somatic mutations detected in lung cancer and molecular tests may enter clinical practice for NSCLC in the near future.
clude mutations in BRAF, HER2(ERBB2), AKT1, MAP2K1 (MEK1), and PIK3CA; gene amplifications in MET; or fusions involving the ROS tyrosine kinase. Each of these is present in a small percentage of patients with lung cancer (1%–5%), and will require further validation before being placed into general use to direct treatment.\textsuperscript{168}

Rebiopsies
A serious limitation in lung cancer molecular testing is availability of tumor tissue. One recommendation to increase the amount of tissue available for testing is to use residual cells from a fine needle aspiration (FNA) to make a paraffin block. This can be done by rinsing the needle into a tissue block. Alternatively, during an FNA or CT-guided core needle biopsy, additional passes can be obtained through the tumor tissue. Patients could also undergo rebiopsy if molecular testing is needed. Notably, the ASCO guidelines also support efforts to increase the availability of tumor tissue for testing.\textsuperscript{169}

Molecular Testing in Prostate Cancer
Table 9 summarizes current molecular biomarkers in prostate cancer.

Prostate-Specific Antigen (\textit{KLK3})
Prostate-specific antigen (PSA) is a peptidase (gene designation, \textit{KLK3}) found in seminal plasma. Serum levels of this protein rise with the presence of prostate cancer and other conditions, such as benign prostatic hyperplasia or inflammation.

Several large-scale screening trials in the United States and Europe\textsuperscript{170–172} have shown a marginal, albeit significant, clinical outcome benefit (decreased deaths from prostate cancer) from use of the serum PSA(KLK3) test with digital rectal examination (DRE) in screening populations of healthy men for prostate cancer. Testing for PSA(KLK3) leads to earlier identification of small tumors, and an increased potential for overtreating tumors that may not be life-threatening. Guidelines for the appropriate use of PSA(KLK3) testing in prostate cancer screening continue to evolve (e.g., NCCN Guidelines for Prostate Cancer Early Detection and for Prostate Cancer; to view the most recent version of these guidelines, visit the NCCN Web site at www. NCCN.org). Changes in PSA(KLK3) level over time, or PSA(KLK3) velocity testing, may also be used to detect disease in patients who have low to intermediate risk of prostate cancer on initial testing,\textsuperscript{173} and thus to detect potentially aggressive tumors while local treatment is still possible.

Serum PSA(KLK3) testing is also used to monitor prostate cancer after diagnosis, and consistently rising levels are used as one indicator of potential disease progression.\textsuperscript{174} After definitive local treatment (e.g., radical prostatectomy or radiation therapy) for clinically localized prostate cancer, serum PSA(KLK3) levels generally drop to undetectable or very low levels. Rising levels of serum PSA(KLK3) in these patients indicates disease recurrence, and monitoring serum PSA(KLK3) levels in these men is of clinical utility.\textsuperscript{175,176}

Table 9 Current Molecular Biomarkers in Prostate Cancer

<table>
<thead>
<tr>
<th>Biomarker</th>
<th>Molecular Compartment</th>
<th>Purpose</th>
<th>Analytic Validity Demonstrated</th>
<th>Level of Evidence</th>
<th>NCCN Category of Evidence</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSA(KLK3)</td>
<td>Serum protein</td>
<td>Diagnostic</td>
<td>Immunoassay</td>
<td>IA</td>
<td>2A</td>
<td>Schroder et al.\textsuperscript{172}</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Hugosson et al.\textsuperscript{171}</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Andriole et al.\textsuperscript{170}</td>
</tr>
<tr>
<td>PCA3</td>
<td>Urine RNA</td>
<td>Diagnostic</td>
<td>RT-PCR</td>
<td>IB</td>
<td></td>
<td>Hessels and Schalken\textsuperscript{178}</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Freedland and Partin\textsuperscript{174}</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Chun et al.\textsuperscript{177}</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Nogueira et al.\textsuperscript{235}</td>
</tr>
<tr>
<td>CTCs</td>
<td>Whole blood</td>
<td>Prognostic</td>
<td>Immunomagnetic separation</td>
<td>IIA</td>
<td></td>
<td>de Bono et al.\textsuperscript{236}</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Danila et al.\textsuperscript{237}</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Scher et al.\textsuperscript{182}</td>
</tr>
</tbody>
</table>

Abbreviations: CTC, circulating tumor cell; RT-PCR, reverse transcriptase polymerase chain reaction.
Although absolute thresholds remain a topic of discussion and controversy, use of PSA(KLK3) measurement in combination with DRE and Gleason score to diagnose prostate cancer is given a category 2A recommendation in the NCCN Guidelines for Prostate Cancer, and can be given a combined level of evidence score of IA owing to long-term, prospectively designed, biomarker-focused clinical studies. Commercial tests for the detection of PSA(KLK3) have been cleared for use by the FDA.

**PCA3**

Although the use of serum PSA(KLK3) testing as a screening tool for prostate cancer is widespread, it has well-known limitations as an early-detection biomarker. Additional biomarkers that can improve on the sensitivity and specificity for early detection and monitoring of prostate cancer continue to be of great interest.

Several molecular alterations have been shown to be common in prostate cancer specimens, which are highly specific to the disease. The most well studied of these, with potential as clinical tools, are DNA methylation and overexpression of certain RNAs in the urine. One of these markers that has shown clinical promise is the PCA3 RNA test. PCA3 (DD3) is a noncoding RNA that can be detected in the urine after DRE, and which is highly expressed in prostate cancer compared with normal prostate. A test for urine PCA3 is commercially available (PROGENSA PCA3 assay, Gen-Probe Inc., San Diego, CA), and the company has submitted a PMA application to the FDA.

Currently, 2 potential clinical uses of this marker are undergoing extensive evaluation. The first (the indication for which the PROGENSA assay is currently being considered by the FDA) is to help determine which of the men with elevated serum PSA(KLK3) and who have already undergone one negative biopsy should undergo a repeat biopsy. The second potential indication is to guide initial biopsy decisions.

In terms of guiding decisions for a repeat biopsy, a large number of studies have been performed, most of which have suggested that the PCA3 test is more specific for predicting a positive repeat biopsy than are serum PSA(KLK3) levels or changes in PSA(KLK3) over time. Furthermore, a recent study was conducted from the placebo arm of the REDUCE trial, which was a prospective randomized clinical trial to determine whether dutasteride was associated with a decreased risk of prostate cancer after 2 or 4 years. In this study, the PCA3 scores were measurable from 1072 of the 1140 subjects, for a 94% informative rate. The use of PCA3 in combination with serum PSA(KLK3) and other risk factors was found to significantly increase diagnostic accuracy in predicting a positive biopsy. These results, using this analytically validated assay, indicate a level of evidence of IB. The clinical utility of PCA3 testing must be interpreted with caution, because the impact of refining decisions when to perform repeat diagnostic biopsies has not been shown.

Several studies have also shown the PCA3 has a high specificity for predicting an initial positive biopsy in men with elevated PSA(KLK3) levels (studies usually conducted in men with serum PSA[KLK3] levels of 4–10 ng/mL). Furthermore, this marker was recently evaluated in a European, prospective, multicenter study in which men with a serum PSA(KLK3) level of 2.5 to 10 ng/mL all underwent prostate needle biopsies. The study was considered a “real life clinical practice study” by the authors and consisted of 516 men. ROC curves showed a significantly higher area under the curve for the PCA3 score versus total PSA(KLK3), PSA(KLK3) density, and percent free PSA(KLK3). Furthermore, the PCA3 score was significantly higher in men with a biopsy Gleason score of 7 or greater versus those with a score less than 7, and in men in whom greater than 33% of the needle cores were positive for cancer versus men with fewer than 33% of cores positive. Overall the inclusion of the PCA3 in multivariable models increased the predictive accuracy by up to 5.5%. These results, therefore, also indicate level IB clinical validity, albeit with a modest improvement in diagnostic accuracy and insufficient evidence to determine the true clinical utility of this assay in routine clinical management.

**CTCs**

The CellSearch CTC detection system has been cleared for use in metastatic prostate cancer by the FDA. However, this test is not included in the NCCN Guidelines for Prostate Cancer (available at www.nccn.org). The test uses an immunomagnetic system to isolate cells with the features of epithelial cells from the peripheral blood and provides an absolute count of CTC per 7.5 mL blood sample. Presence of greater than 5 CTCs per 7.5 mL sample at baseline measurement has been shown to
be associated with shortened overall survival, an effect which is seen through the course of several CTC measurements. The association with elevated levels of CTCs and poor prognosis is consistent and strong, and therefore the level of evidence is IIB. In addition, change in numbers of CTCs in response to therapy has recently met the Prentice criterion of surrogacy for overall survival in the setting of treatment with abiraterone. Thus, decreased CTC levels from greater than 5 to lower than 5 is a predictive biomarker, although it is only informative once therapy has begun, with a level of evidence of IIa. Currently, the clinical utility of changing CTC levels remains untested, because no trial has evaluated whether changing therapy based on a lack of CTC response improves outcome.

Similar to other forms of cancer, comprehensive molecular analyses of prostate cancer using microarray and sequencing technologies have resulted in a greater understanding of tumor biology. It remains clear that the androgen receptor is a driver of both treatment-naive and castration-resistant prostate cancer, but biomarkers that accurately reflect androgen receptor activity have yet to be developed. Similarly, most prostate cancers contain fusion proteins between androgen-regulated genes and members of the ETS family of transcription factors. These fusion genes are likely also oncogenic drivers, but data conflict regarding their prognostic strength as biomarkers.

Molecular Testing in Acute Myeloid Leukemia

Acute myeloid leukemia (AML) is a genetically and clinically heterogeneous neoplasm characterized by the accumulation of genetic and epigenetic alterations in hematopoietic progenitor cells that alter normal mechanisms of differentiation, proliferation, and survival. Cytogenetics have been used for outcome prediction and treatment guidance, because they were shown to enable patients with AML to be stratified into cytogenetic risk groups (i.e., favorable, intermediate, and unfavorable) based on the presence of nonrandom chromosome aberrations or normal karyotype. However, within each cytogenic group, it has become clear that molecular heterogeneity exists, which can be further used to determine molecular risk stratification and treatment selection.

Few of the genetic molecular alterations have shown relevance to the biologic and clinical classification of AML as first recognized by the WHO classification, and more recently have been incorporated into clinical guidelines such as the NCCN Guidelines for AML and the European LeukemiaNet guidelines. Consensus exists regarding clinical impact for at least for 4 of the markers (NPM1, CEBPA, and KIT mutations, and FLT3 internal tandem duplication [ITD]). However, the best therapeutic approaches for patients harboring these mutations remain to be defined. Table 10 summarizes current molecular biomarkers in AML.

NPM1

Mutations in exon 12 of the NPM1 gene are found in approximately one-third of adult patients with AML, and result in abnormal cytoplasmic localization of the NPM1 protein. NPM1 mutations are frequent in cytogenetically normal AML (CN-AML) but can also be found in other cytogenetic or molecular groups, most frequently in patients harboring FLT3 mutations. NPM1 mutations, particularly the genotype “mutated NPM1 without concurrent FLT3-ITD,” has been associated with achievement of complete remission and favorable outcome. Recent data also suggest that NPM1 mutations seem to predict better outcome in older patients (> 60 years of age), and perhaps could be used to identify individuals in this age group who may benefit from intensive conventional chemotherapy. However, these results require validation.

CEBPA

Approximately 15% of patients with CN-AML harbor CEBPA mutations. These mutations, which may also be found in other cytogenetic groups of AML, usually affect the N-terminal region of the gene, thereby resulting in a mutant protein similar to a truncated CEBPA isoform found also in normal cells or the C-terminal basic region and leucine zipper domain of the gene. Both mutations may result in a mutant protein with altered transcription factor activity with respect to the wild-type counterpart. In approximately two-thirds of cases, C- and N-terminal mutations are “biallelic” mutations (also called double mutations), with most (~90%) being compound heterozygous (C-terminal on one allele and N-terminal on the other) and the rest being homo-
### Table 10 Current Molecular Biomarkers in Acute Myeloid Leukemia

<table>
<thead>
<tr>
<th>Biomarker</th>
<th>Molecular Compartment</th>
<th>Purpose</th>
<th>Analytic Validity Demonstrated</th>
<th>Level of Evidence</th>
<th>NCCN Category of Evidence</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Markers With Accepted Clinical Utility</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FLT3-ITD</td>
<td>Tumor DNA</td>
<td>Predictive/prognostic:FLT3-ITD mutation confers poor risk status</td>
<td>Multiple methods: PCR, multiplex assays, direct sequencing</td>
<td>IIB</td>
<td>Prognostic: 2A</td>
<td>Schlenk et al.193</td>
</tr>
<tr>
<td>CEBPA mutation</td>
<td>Tumor DNA</td>
<td>Predictive/prognostic: CEBPA mutation with normal cytogenetics and in absence of FLT3-ITD confers better risk status</td>
<td>Multiple methods: PCR, multiplex assays, direct sequencing</td>
<td>IIB</td>
<td>Prognostic: 2A</td>
<td>Marcucci et al.197, Burnett et al.194, Taskesen et al.198, Wouters et al.199</td>
</tr>
<tr>
<td>NPM1 mutation</td>
<td>Tumor DNA</td>
<td>Predictive/prognostic: NPM1 mutation with normal cytogenetics and in absence of FLT3-ITD confers better risk status</td>
<td>Multiple methods: PCR, multiplex assays, direct sequencing</td>
<td>IIB</td>
<td>Prognostic: 2A</td>
<td>Becker et al.194</td>
</tr>
<tr>
<td>KIT mutation</td>
<td>Tumor DNA</td>
<td>Predictive/prognostic: c-KIT mutations in the presence of t(8;21), inv(16), or t(16;16) confers a higher risk of relapse</td>
<td>Multiple methods: PCR, multiplex assays, direct sequencing</td>
<td>IIB</td>
<td>Prognostic: 2A</td>
<td>Paschka200, Paschka et al.201</td>
</tr>
<tr>
<td><strong>Markers With Emerging Evidence</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FLT3-TKD mutation</td>
<td>Tumor DNA</td>
<td>Prognostic relevance of FLT3-TKD mutations is controversial</td>
<td>Multiple methods: PCR, multiplex assays, direct sequencing</td>
<td>Insufficient data</td>
<td></td>
<td>Schlenk et al.193</td>
</tr>
<tr>
<td>WT1 mutation</td>
<td>Tumor DNA</td>
<td>WT1 mutations associated with poorer outcome</td>
<td>Multiple methods: PCR, multiplex assays, direct sequencing</td>
<td>IIB</td>
<td></td>
<td>Paschka et al.205, Virappane et al.206, Gaidzik et al.204</td>
</tr>
<tr>
<td>RUNX1 mutation</td>
<td>Tumor DNA</td>
<td>RUNX1 mutations associated with poorer outcome</td>
<td>Multiple methods: PCR, multiplex assays, direct sequencing</td>
<td>IIB</td>
<td></td>
<td>Tang et al.203</td>
</tr>
<tr>
<td>MLL-PTD</td>
<td>Tumor DNA</td>
<td>MLL-PTD associated with inferior CR duration and relapse free survival</td>
<td>RT-PCR</td>
<td>IIB</td>
<td></td>
<td>Whitman et al.202</td>
</tr>
<tr>
<td>IDH1 mutation</td>
<td>Tumor DNA</td>
<td>IDH1 mutations associated with NPM1 mutation and predict worst outcome for patients with mutated NPM1 without FLT3-ITD</td>
<td>Multiple methods: PCR, multiplex assays, direct sequencing</td>
<td>IIB</td>
<td></td>
<td>Mardis et al.187, Marcucci et al.207, Paschka et al.208</td>
</tr>
<tr>
<td>IDH2 c.515G&gt;A p. R172K</td>
<td>Tumor DNA</td>
<td>IDH2 R172 may confer lower probability of achieving CR and possibly also inferior outcome</td>
<td>Multiple methods: PCR, multiplex assays, direct sequencing</td>
<td>IIB</td>
<td></td>
<td>Mardis et al.187, Marcucci et al.207, Paschka et al.208</td>
</tr>
<tr>
<td>IDH2 codon 140 mutation</td>
<td>Tumor DNA</td>
<td>prognostic relevance of IDH2 codon 140 mutation is controversial</td>
<td>Multiple methods: PCR, multiplex assays, direct sequencing</td>
<td>IIB</td>
<td></td>
<td>Mardis et al.187, Marcucci et al.207, Paschka et al.208</td>
</tr>
</tbody>
</table>

Abbreviations: CR, complete remission; ITD, internal tandem duplication; PCR, polymerase chain reaction; PTD, partial tandem duplication; RT-PCR, reverse transcriptase polymerase chain reaction; TKD, tyrosine kinase domain.
zygous (either both C- or N-terminal).\textsuperscript{196–198} CEBPA mutations have consistently been associated with a relatively favorable outcome for patients with CN-AML.\textsuperscript{196–198} However, more recent studies indicate that the favorable impact is restricted to patients with double mutations.\textsuperscript{199}

Patients with CEBPA mutations are grouped together with those with AML with mutated NPM1 without FLT3-ITD and those with core binding factor (CBF) AML [t(8;21) or inv(16)/t(16;16)] in a favorable-risk group.\textsuperscript{191} Therapeutic recommendations for these subsets are similar and consist of chemotherapy-based approaches, whereas allogeneic hematopoietic stem cell transplantation (HSCT) should be reserved for only patients who have experienced relapse or have proven additionally adverse cytogenetic or molecular features.

**FLT3**

Two types of mutations that result in the constitutive activation of the encoded tyrosine kinase receptor protein have been identified in the FLT3 gene: tandem internal duplications within the juxtamembrane domain and point mutations with the tyrosine kinase domain (TKD). The FLT3-ITD is found in approximately 20% of unselected patients with AML and approximately 30% of those with CN-AML.\textsuperscript{193} The prognosis of patients with CN-AML with FLT3-ITD is significantly inferior compared with CN-AML without the mutation when treated with current standard chemotherapy, particularly if the ratio of mutated versus wild-type alleles is high.\textsuperscript{193} The prognostic relevance of FLT3-TKD mutations instead is controversial. Although patients with FLT3-ITD have been reported to benefit from allogeneic HSCT, this approach must be validated prospectively. Furthermore, because FLT3-ITD encodes a constitutively activated tyrosine kinase protein, small inhibitor molecules combined with chemotherapy are being tested in phase III clinical trials.

**KIT**

KIT mutations are found in 25% to 30% of CBF-AML and are rare in other AML subsets.\textsuperscript{200} In most studies, KIT mutations have been associated with inferior outcome.\textsuperscript{201} Notably, KIT is not only mutated but also expressed at significantly higher levels in CBF-AML compared with other AML subsets. Currently no data support the use of KIT mutational status to guide therapy, although these patients may benefit from more-aggressive treatment with allogeneic HSCT. Clinical trials are currently underway evaluating KIT inhibitors in CBF-AML.

For use as prognostic markers, FLT3-ITD, KIT, CEBPA, and NPM1 mutations are all given an NCCN category of 2A. As predictive markers, all have an NCCN category of 2B. Multiple large-scale retrospective studies of marker expression and outcome for these mutations gives them a combined level of evidence of IIB.

**Other Gene Mutations**

In addition to the aforementioned mutations, other genetic mutations have been found in patients with AML. These mutations either do not seem to significantly contribute to risk stratification or have so far been less well studied for their prognostic significance.

**MLL:** Partial tandem duplications (PTD) of MLL are found in 5% to 11% of patients with CN-AML and frequently in AML with trisomy 11. MLL-PTD have been associated with inferior complete remission duration and relapse-free survival, although more recent studies show no prognostic impact in patients with CN-AML intensively treated with autologous HSCT or 4 cycles of chemotherapy consolidations.\textsuperscript{202}

**RUNX1:** RUNX1 is deregulated in AML by chromosomal translocations and by mutations clustering in the Runt domain of the gene. RUNX1 mutations have been associated with undifferentiated (M0) morphology, and with specific chromosome aberrations, such as trisomy 21 and trisomy 13. In recent studies, RUNX1 mutations were found in approximately 13% of patients with AML and associated with worse outcome.\textsuperscript{203}

**WT1:** WT1 mutations are found in 10% to 13% of patients with CN-AML. In most studies, WT1 mutations have been associated with inferior outcome.\textsuperscript{204–206}

**IDH1/IDH2:** Mutations of IDH1 and IDH2 were first reported in gliomas, and only more recently in AML.\textsuperscript{187} The aggregate frequency of IDH1 and IDH2 mutations in AML is relatively high, with approximately 15% to 20% of all patients with AML and 25% to 30% of patients with CN-AML harboring either IDH1 or IDH2 mutations.\textsuperscript{207,208} Initial studies from larger and homogeneous cohorts of patients indicate that IDH1 mutations are significantly associated with NPM1 mutations and predict worse outcome for patients with mutated NPM1 without...
FLT3-ITD. The distinct R172 IDH2 mutation is rarely associated with any of the other known prognostic mutations, and seems to confer lower probability of achieving complete remission and possibly also inferior outcome. The prognostic impact of IDH2 mutation in codon 140 remains instead controversial.

Testing for mutations in these genes is not included in the NCCN Guidelines for AML (available at www.NCCN.org). However, results of large-scale retrospective analyses in samples from clinical trials gives each of them a combined level of evidence score of IIB. Clinical utility remains to be determined.

NRAS, TP53, TET2, DNMT3A, and ASXL1: NRAS, TP53, TET2, DNMT3A, and ASXL1 mutations have also been found in patients with AML, but their clinical impact as prognosticators or predictors of response to distinct treatment approaches remains to be fully defined. These mutations, which have correspondingly fewer data indicating their prognostic or predictive role, are all assigned a combined level of evidence of IIIC for clinical validity.

A continuing challenge in validating biomarkers in AML, as each new marker further divides patients into smaller phenotypic groups, will be to find sufficient numbers of patients to establish the clinical utility of each marker and to discover the impact that each may have on treatment or prognosis.

Conclusions
Cancer care is becoming increasingly dependent on tumor markers to diagnose, anticipate prognosis, and select optimal therapy for patients. Although knowledge of the specific tumor markers that have the greatest value in care will evolve over time, it is important to develop a common language with respect to the specific purpose of each tumor marker, the steps in developing a tumor marker, and the different levels of evidence supporting the specific use of a tumor marker. The task force was established to review each of these processes and to identify the tumor markers that are most pertinent today.

This report presents tumor markers for 6 cancer types and highlights some of the challenges faced by oncologists in determining which tests will provide the most useful information to direct patient care. First, although significant efforts have been made to standardize the development and reporting of diagnostic, prognostic, and predictive tumor markers, best practices have not been broadly adopted and the literature supporting the use of tumor markers is often incomplete or reported in a manner that hampers comparisons across studies. Second, the actual developmental process for tumor markers and the regulatory oversight of that process is relatively immature compared with drug development and, as a result, confusion often occurs regarding the steps required to incorporate tumor markers into care. Third, several technical methods are often available to measure a specific molecular marker, usually with concordance between assays but each with distinct analytic performance characteristics and few opportunities for rigorous head-to-head comparison. Fourth, great interest has been shown among all stakeholders in cancer care, including clinical and basic investigators, providers, and patients, to rapidly incorporate emerging biologic insights into clinical care that drives a desire for early adoption, often before rigorous assessment of the analytic validity, clinical validity, and clinical utility of a tumor marker.

With these challenges in mind, this report focuses on presenting the different uses for tumor markers, an overview of the different steps of tumor marker development, and an adapted, simplified categorization of level of evidence. After reviewing the work previously reported on tumor markers, the task force believes that identifying the pertinent references and efforts previously focused on biomarker standardization and classification of levels of evidence for biomarkers is an important first step. Admittedly, challenges remain with respect to the development of a simple yet comprehensive classification system for levels of evidence. The current systems focus only on the supportive data and not on panel consensus, as in the NCCN Guidelines. With time, the classification of level of evidence for tumor markers and the common accepted criteria will likely be increasingly refined.

Looking forward, the task force believes that the community interested in developing tumor markers should consider and adopt the following recommendations:

- Assure a test is analytically valid before performing rigorous testing for clinical validity.
- If the test is a qualitative one, such as immunohistochemistry, determine the degree to which the test has been rigorously quality controlled.
NCCN Task Force Report

Have preferred methods or protocols been published?
- When evaluating new tests, check to see if BRISQ standards or similar have been used for preanalytic sample collection in those trials.
- When evaluating new studies, check to see if REMARK standards were used in clinical trial design for marker study and reporting of the results.
- Consider whether the test as it currently exists can be scaled up for general use, is an established LDT in a CLIA-certified laboratory, or is an FDA-approved/cleared test.
- Include comparisons of testing methods and make recommendations for preferred methods.
- Include and consider recommendations from other groups, including CAP, ASCO, and EGAPP. If recommendations/outcomes differ, address them specifically so that clinicians can understand the source of potential disagreement, and thus improve decision-making.
- Incorporate a formal discussion of levels of evidence (Tables 1–3) to determine clinical utility for tests under consideration.
- Include a discussion of needed studies or evidence for tests under consideration if gaps are perceived in the literature.

The use of tumor markers will continue to transform cancer care. Only through adopting a common language and standards of validation will timely and definitive studies on emerging markers be effectively communicated, thereby ensuring that clinicians take full advantage of the current genomic era.

Acknowledgments

The authors would like to acknowledge Jason T. Huse, MD, PhD, Cyrus V. Hedvat, MD, PhD, and a molecular testing expert from the Centers for Medicare & Medicaid Services (CMS) for critical reading of sections of this manuscript.

Representatives from the FDA and the NCI served as ex officio task force members.

References:
23. Edge SB, Byrd DR, Compton CC, et al. AJCC Cancer Staging
Tumor Markers in Oncology


Tumor Markers in Oncology


## Individual Disclosures for NCCN Task Force: Evaluating the Clinical Utility of Tumor Markers in Oncology Panel Members

<table>
<thead>
<tr>
<th>Panel Member</th>
<th>Clinical Research Support</th>
<th>Advisory Boards, Speakers Bureau, Expert Witness, or Consultant</th>
<th>Patent, Equity, or Royalty</th>
<th>Other</th>
<th>Date Completed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phillip G. Febbo, MD</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>5/12/11</td>
</tr>
<tr>
<td>Marc Ladanyi, MD</td>
<td>None</td>
<td>Infinity Pharmaceuticals</td>
<td>None</td>
<td>None</td>
<td>5/12/11</td>
</tr>
<tr>
<td>Kenneth D. Aldape, MD</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>5/10/11</td>
</tr>
<tr>
<td>Robert L. Becker, Jr., MD, PhD</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>5/9/11</td>
</tr>
<tr>
<td>Marian Birkeland, PhD</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>GlaxoSmithKline</td>
<td>5/7/11</td>
</tr>
<tr>
<td>Barbara Conley, MD</td>
<td>sanofi-aventis; Endocyte, Inc.; and ImClone Systems</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>5/9/11</td>
</tr>
<tr>
<td>Angelo M. De Marzo, MD, PhD</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>5/12/11</td>
</tr>
<tr>
<td>M. Elizabeth Hammond, MD</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>5/7/11</td>
</tr>
<tr>
<td>Daniel F. Hayes, MD</td>
<td>GlaxoSmithKline; Pfizer Inc.; Novartis AG; and Veridex, LLC</td>
<td>Chugai Pharmaceutical Co., Ltd; and BioMarker Strategies</td>
<td>Immunicon Corporation; University of Michigan; and Oncimmune LLC</td>
<td>None</td>
<td>5/10/11</td>
</tr>
<tr>
<td>A. John Iafrate, MD, PhD</td>
<td>None</td>
<td>Pfizer Inc.; and Abbott Laboratories</td>
<td>None</td>
<td>None</td>
<td>5/10/11</td>
</tr>
<tr>
<td>R. Kate Kelley, MD</td>
<td>Genomic Health, Inc.</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>5/11/11</td>
</tr>
<tr>
<td>Guido Marcucci, MD</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>5/12/11</td>
</tr>
<tr>
<td>Shuji Ogino, MD, PhD, MS</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>5/13/11</td>
</tr>
<tr>
<td>William Pao, MD, PhD</td>
<td>Enzon Pharmaceuticals, Inc.; AstraZeneca Pharmaceuticals LP; and Xcovery</td>
<td>Clovis Oncology; Bristol-Myers Squibb Company; AstraZeneca Pharmaceuticals LP; MolecularMD; and Symphony Evolution, Inc.</td>
<td>MolecularMD</td>
<td>None</td>
<td>5/13/11</td>
</tr>
<tr>
<td>Dennis C. Sgroi, MD</td>
<td>None</td>
<td>bioTheranostics, Inc.</td>
<td>None</td>
<td>None</td>
<td>5/12/11</td>
</tr>
</tbody>
</table>
Post-test

1. In order for molecular test results to be used to determine patient care, the test must be:
   a. FDA-approved
   b. FDA-cleared
   c. Performed in a CLIA laboratory
   d. Any of the above
   e. All of the above

2. If a given test is approved or cleared by the FDA, this indicates that it has proven clinical utility. Is this statement true or false?
   a. True
   b. False

3. What is detected by the companion diagnostic recently FDA approved with vemurafenib for treatment of metastatic melanoma?
   a. ALK gene fusion
   b. ESR1 overexpression
   c. BRAF V600 mutation
   d. KRAS mutation
   e. Serum PSA

4. A hypothetical test for determining somatic mutation in a newly identified gene has been shown to predict response to a new chemotherapy agent. If use of the test is required for prescription of the drug, and both are approved by the FDA, the test would be called a:
   a. Laboratory-developed test (LDT)
   b. CLIA-certified test
   c. Companion diagnostic
   d. CAP-certified test

5. Which of the following must have proven clinical validity?
   a. An LDT performed in a CLIA laboratory
   b. An FDA-approved test
   c. An FDA-cleared test
   d. A molecular test performed in a research laboratory
   e. a and b
   f. b and c
   g. a, b, c, and d
   h. a and d
   i. None of the above

6. The Clinical Laboratory Improvement Amendment (CLIA) of 1988 is administered by:
   a. FDA
   b. CMS
   c. CAP
   d. NCCN

7. What can be done to help prove clinical utility of a molecular test?
   a. Demonstrate that test performance is reproducible in 3 or more clinical laboratories.
   b. Demonstrate that test results predict response/lack of response to a given therapy.
   c. Demonstrate that test results show which subset of patients have the longest progression-free survival.
   d. Demonstrate that test results are statistically significant.

8. Which molecular test is predictive for response to trastuzumab?
   a. HER2(ERBB2) protein expression by immunohistochemistry
   b. ER-α(ESR1)/PgR(PR) protein expression by immunohistochemistry
   c. HER2(ERBB2) amplification by FISH
   d. a and b
   e. a and c
   f. a, b, and c

9. Which tests have proven clinical utility in NSCLC?
   a. EGFR mutation and KRAS mutation
   b. ERCC1 expression and BRAF mutation
   c. ALK gene fusion and EGFR mutation
   d. ALK gene fusion and KRAS mutation

10. Which molecular tests have proven clinical utility and outperform Gleason sum as a prognostic for localized prostate cancer?
    a. Baseline PSA level
    b. Urine PCA3 testing
    c. Circulating tumor cell analysis
    d. Gene expression analysis
    e. All of the above
    f. None of the above

11. IDH1 and IDH2 mutations have been detected in what types of cancers:
     a. AML
     b. NSCLC
     c. Glioma
     d. a and b
     e. a and c
     f. b and c
     g. None of the above

12. Which methods of testing have proven analytic validity for MSI/MMR assessment in colon cancer?
    a. PCR for MSI
    b. Flow cytometry for MSI
    c. Immunohistochemistry for MMR
    d. FISH for MMR
    e. a and b
    f. a and c
    g. b and c
    h. c and d

To Receive Credit

To receive credit, participants will read all portions of this monograph, including all tables, figures, and references. To receive your continuing education credit and certificate, visit http://www.cvent.com/d/rdq8rk to complete the post-test and evaluation.

All post-test scores must be received by midnight on November 25, 2012, in order to be eligible for credit.

It should take approximately 0.75 hours (45 minutes) to complete the activity as designed. There is no registration fee for this activity.