Barriers to Integrating Gene Profiling for Stage II Colon Cancer
Scott Kopetz and James L. Abbruzzese

In this issue of Clinical Cancer Research, Jorissen and colleagues provide evidence that gene expression changes can be used to predict outcome for early stage colorectal cancers, thereby providing a useful addition to the extensive studies evaluating risk factors for recurrence of this disease (1). The clinical value of such risk factors derives from the ability to direct chemotherapy appropriately and withhold it from patients unlikely to benefit from treatment. As such, stage II (and Dukes B) colon patients are the subset of patients most likely to benefit from these gene signatures as there remains equipoise about the utility of adjuvant chemotherapy in this group.

Currently, there are several accepted pathologic features associated with an increased risk of recurrence for stage II disease. Tumor (T-stage) and nodal stage (N-stage) have been known to be the most important features and form the basis of the widely used TNM staging system. Recent studies have shown that the number of nodes resected and examined is an independent risk factor for recurrence, likely demonstrating an impact of quality of surgical resection, thoroughness of pathology review, and perhaps an underlying impact of patient biology (2). Other recognized poor prognosis factors include lymphovascular invasion and poorly differentiated tumors. Current guidelines incorporate these features into a stage II phenotype in which it is anticipated that adjuvant chemotherapy can provide increased benefit (3).

However, despite these clinical features, significant variability remains in prognosis for stage II and III patients. Gene expression profiling using microarrays has rapidly evolved as a potential approach to reduce the complexity of both known as well as previously unidentified risk factors to a single score or category. Thus, several groups have used gene profiling techniques to better define outcomes for these patients, primarily by using retrospective samples variably treated with adjuvant chemotherapy (4–9).

The study in this issue uses a different methodology to derive a high-risk prognostic score (1). Prognostic markers are classically defined as determinants of outcome independent of treatment. As a result, bias is introduced if these markers are derived from a set of patients who have been heterogeneously treated or only from a subset of the population irrespective of treatment. By profiling tumors from early (stage I) and late (stage IV) primary tumors, the authors have devised a scoring system that can be applied to stage II and III patients. With this approach, they were able to successfully classify the vast majority of stage II and III patients into stage I- or stage IV-like tumors. Those stage III tumors with stage IV-like profiles had a 2.9 fold increased hazard of recurrence than those patients with stage I-like profiles. Similarly, stage II tumors with a stage IV-like profile had a 10.3 hazard ratio for recurrence compared with stage I-like tumor.

Importantly, further development of gene expression-based prognostic scores could be strengthened by increased incorporation of the widely accepted pathologic high-risk features into the prognostic models. It remains to be seen if these prognostic scores duplicate these clinical features or provide independent information on clinical outcome. In the multivariate model presented by Jorissen and colleagues, several clinical features were included, with a hazard ratio of 8.6 for the gene profile in stage II patients. However, the T-stage, the strongest clinical predictor of outcome for stage II patients, was not included because of the limited sample size. Similarly, when the number of lymph nodes involved with cancer (N-stage) was included in the multivariate analysis for stage III patients, the gene profile was no longer significant.

Genetic features of colon cancer are increasingly being incorporated into clinical care and should be included in assessment of microarray profiles. For example, studies have reproducibly shown the prognostic significance of microsatellite instability, a hallmark of hereditary nonpolyposis colon cancer (or Lynch syndrome), which is also found in a subset of sporadic colon cancers. Stage II patients with microsatellite instability have dramatically better clinical outcomes. Development of colon cancer in the presence of microsatellite instability seems to be independent of the classic chromosomal instability pathway. It is, therefore, certainly plausible that the gene expression profiles may be only recapitulating these disparate biologies and outcomes.

Although clinicians commonly use prognostic characteristics to assign treatment to stage II patients, what is ultimately
needed are predictive profiles that can identify those stage II patients who will benefit from adjuvant therapy. The assumption that higher risk subsets of patients will derive more absolute benefit from therapy may not hold for all molecular subtypes of colon cancer. As a cautionary example, despite the fact that stage II patients with microsatellite instability have excellent outcomes without adjuvant therapy, they seem to fare worse when treated with 5-fluorouracil adjuvant chemotherapy. Unfortunately, predictive biomarkers are much more difficult to validate, as they classically require a cohort of patients of equal risk randomly assigned to treatment or a control without chemotherapy.

There are several barriers to implementation of these gene profiles. Although separate validation sets are now common for most published studies, external and independent validation remains a critical step. However, given the costs and limited patient resources, it is rare for this next validation step to be done. An increasingly recognized approach is a meta-analysis of microarray data. Meta-analyses that compare the prognostic genes in published articles can identify critical genes that may account for the prognostic effect or over-representation of genes from a common critical pathway. A standardized format for collection of raw microarray data is in use, but a parallel format for standardized reporting of results remains to be widely implemented and limits the meta-analytic assessment of gene lists (10). Such a review of the gene list of Jorissen and colleagues shows very little overlap with previously published lists (Fig. 1). These apparently disparate results may reflect identification of the same higher risk subtypes using different, but closely correlated genes, but proof of such a relationship requires access to per-patient expression data and outcomes. This second type of meta-analysis, which uses per-patient data instead of gene-lists, has the potential to be a practical and low-cost method of external validation. The challenges of comparing across studies and microarray platforms have limited the initial impact of this approach, although bioinformatics tools to assist in this endeavor have been developed (11).

Other novel approaches are needed for such research of prognostic gene profiles. The cost of these studies, although declining, still remains high. This forces smaller sample sizes and limits power. Even in this moderately sized study, the confidence intervals on the hazard ratios remain large (95% confidence interval of 1.3 to 80.0 for high-risk stage II tumors). Novel approaches such as sequential analysis, loosely analogous to a two-stage design of clinical trials, may identify strongly positive or clearly uninformative studies sooner, thereby limiting resource expenditure (12).

Ultimately, our success in treating early stage colon cancer is limited by the effectiveness of our chemotherapy and the ability to integrate molecular and clinical data efficiently in the clinic. For now, the extraordinary promise of a molecular profile for early stage colon cancer seems to be equally matched by the hurdles to appropriate validation and application of these findings to the clinic.

**Disclosure of Potential Conflicts of Interest**

No potential conflicts of interest were disclosed.

**References**

Clinical Cancer Research

Barriers to Integrating Gene Profiling for Stage II Colon Cancer

Scott Kopetz and James L. Abbruzzese

Clin Cancer Res  Published OnlineFirst December 8, 2009.

Updated version  Access the most recent version of this article at: doi:10.1158/1078-0432.CCR-09-2523
Supplementary Material  Access the most recent supplemental material at: http://clincancerres.aacrjournals.org/content/suppl/2009/12/15/1078-0432.CCR-09-2523.DC1

E-mail alerts  Sign up to receive free email-alerts related to this article or journal.
Reprints and Subscriptions  To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.
Permissions  To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.