Molecular Markers for Early Detection of Renal Carcinoma: Investigative Approach

Steven Skates and Othon Iliopoulos

1Biostatistics Center, Department of Medicine, Massachusetts General Hospital, Boston, Massachusetts, and 2Hematology-Oncology Unit, Department of Medicine, Massachusetts General Hospital Cancer Center, Boston, Massachusetts

ABSTRACT

Molecular markers for renal cell carcinoma could guide early detection of localized disease in defined populations at high risk for the disease or early disease recurrence after nephrectomy for renal cell carcinoma. In addition, sensitive and specific markers may provide surrogate end points for clinical trials of treatment and/or disease prevention. Powerful techniques of genomic and proteomic analysis of human renal carcinoma cell lines, tumor samples, and biological fluids, such as plasma and urine, obtained from patients with renal cell carcinoma, are likely to identify candidate markers. Careful selection among early candidate markers and further testing in independent patient populations are required for marker validation.

INTRODUCTION

The incidence of renal cell carcinoma in the United States has been estimated at approximately 31,000 cases per year. This incidence has been rising since 1970 for both men and women and whites and blacks. The Surveillance Epidemiology and End Results program estimates that 31,900 new cases of renal cell carcinoma will be diagnosed in 2003, bringing the number of estimated death from renal cell carcinoma to 11,900 this year. For comparison, the estimated incidence for ovarian cancer is 25,000 and for brain tumors is 18,000. The highest incidence is among American blacks older than 65 years (82 cases per 100,000 population; ref. 1).

Approximately 30% of patients with kidney cancer present with metastatic disease at the time of diagnosis. Metastatic disease patients have a 5-year survival rate of 5% compared with 50 to 95% in patients with localized disease (2). Thus, if it were feasible to develop a screening test to detect early kidney cancer, mortality associated with the disease would likely be substantially reduced.

This straightforward view of benefits from renal cell carcinoma biomarker(s) meets with a formidable statistical challenge presented by the relative low prevalence of the disease. Even if the biomarker for renal cell carcinoma was 100% sensitive and had a specificity of 99.4%, the positive predictive value for men older than 65 years would be only 10%. Achieving such an exacting combination of sensitivity and specificity is unlikely for a blood test, even using multiple markers and/or longitudinal measurements (3). A high specificity of 99.9% and a reasonable sensitivity of 70% was achieved for ovarian cancer with a multimodal approach to screening where a blood test was used as a first-line test, and if elevated, the subject was referred to an imaging test, in this case, transvaginal sonography (4). The first-line test achieved a specificity of 98%, and the second line test increased it to 99.9%, giving a positive predictive value of 20%, doubling the minimum positive predictive value of 10% considered acceptable (5). Such a multimodal approach is likely to be used for screening renal cell carcinoma, with follow-up imaging either being an ultrasound or a computed tomography scan, and hence, the requirements for the first-line blood test are likely to be similar, namely a specificity near 98% and a high sensitivity. For a high-risk population, where the incidence is increased by an order of magnitude (typically 10-fold), the specificity requirement to achieve the same positive predictive value is lower. Thus, identifying and applying early candidate biomarkers in populations at high risk for renal cell carcinoma is an efficient route for determining likely candidate screening markers for identifying subjects in the normal risk population at sufficient risk for referral to the more expensive imaging modality as a second-line test.

Similarly, patients who have previously been diagnosed and treated for renal cell carcinoma have a higher incidence of recurrence, and for this reason, the specificity and sensitivity requirements are not as stringent as for screening for new disease. This population also provides an efficient approach to identifying biomarkers for renal cell carcinoma. Although such biomarkers are usually referred to as tumor markers to distinguish them from screening tests, new tumor markers may also be useful as candidate screening markers.

In this article, we review the populations at risk for renal cell carcinoma, examine the biological material available for biomarker interrogation, and summarize the techniques of such an interrogation. Finally, we propose an algorithm of selection among early candidate biomarkers for additional validation.

PATIENTS AT RISK FOR RENAL CELL CARCINOMA

Patients at high risk for developing renal cell carcinoma can be categorized into two broad groups. The first group consists of patients with sporadic renal cell carcinoma who are at high risk for relapse after initial treatment. The second group includes normal individuals carrying a germ-line mutation in one of the genes known to confer a lifetime high risk for
developing renal cell carcinoma. Both groups, described below, can provide target populations for discovery and early validation of biomarkers.

Patients who present with localized, sporadic renal cell carcinoma have an estimated 10 to 50% chance of local relapse or metastasis after nephrectomy (6). This risk largely relates to disease stage, with patients with stage I disease running a risk of 10% for relapse and patients with stage II to III disease having a higher risk of relapse with estimates ranging from 40 to 70% (2).

Patients with renal cell carcinoma face long-term risks for developing a separate renal cell carcinoma in the contralateral kidney. White men and women have relative risks of 13.5 and 5.1, respectively, even after 10 years of follow-up. Black men are at a 40-fold higher risk than an age- and race-matched population without prior renal cell carcinoma for developing contralateral renal cell carcinoma (7).

Defined germ-line mutations in tumor suppressor genes or oncogenes lead to familial predisposition for renal cell carcinoma. Specifically, patients with inactivating mutations in the von Hippel-Lindau, Birt-Hogg-Dube, or fumarate hydratase tumor suppressor genes have a lifetime risk of developing correspondingly clear cell, oncocytic/chromophobe, or papillary type 2 kidney cancer (8–10). In addition, activating germ-line mutations in the c-Met proto-oncogene predispose carriers to papillary type 1 renal cancer (11). In addition, patients undergoing long-term hemodialysis are at higher risk of developing papillary tumors (12).

**TYPE OF BIOMARKERS**

A simplified stratification of biomarkers may recognize their role at different stages of disease development and/or progression (Fig. 1). Schematically, one can assume that for a given clinical presentation of malignancy, there is a latent period of disease development (preclinical phase), a specific time at which diagnosis is made and treatment initiated, followed by a variable period of observation to diagnose and treat any local recurrence or distant relapse or metastasis (follow-up period). Biomarkers can serve several different functions. They can identify individuals at risk for developing the disease (markers of risk), detect individuals with early signs of disease or recurrence (early detection markers), or predict tumor biology or response to therapy (prognostic or selection markers).

Molecular biomarkers for risk are detectable changes, systemic or local, that indicate the carrier of the biomarker is at higher risk for developing the disease in the future—in this case the malignancy. For example, the presence of human papillomavirus at the cervix indicates a higher risk for subsequently developing cervical cancer. Other such examples include the association between Helicobacter pylori and gastric cancer and EBV and nasopharyngeal carcinoma. These examples are dichotomous risk factors in infectious conditions, and although it is conceptually possible for protein markers to be markers of risk, thus far to our knowledge no continuously quantifiable markers have been identified as such.

Although risk markers may identify individuals who warrant greater surveillance than usual, the identification of an individual at high risk does not by itself help with early disease diagnosis. Molecular biomarkers of early detection may allow identification of patients who have already developed clinically detectable disease yet are still in early stages of the disease and more usefully before clinical symptoms occurring. For example, significant increases in prostate-specific antigen or CA-125 levels may indicate the presence of clinically detectable prostate or ovarian cancer, respectively. Because such a marker may correspond to overall disease burden, it may also serve as a surveillance marker for early diagnosis of disease recurrence in the posttreatment follow-up period, as is the case with prostate-specific antigen and CA-125.

Several clinical and laboratory values may serve as markers of overall prognosis and/or predictors of response to therapy at the time of diagnosis. Among these values, often with independent prognostic or predictive value, are molecular markers identified in the tumor specimen (e.g., expression of Her2/neu in breast cancer) or the biological fluids of the host (lactate dehydrogenase as prognostic factor in large cell lymphoma). Gene expression profiling of tumors constitutes such an effort to globally describe the biological behavior of histologically similar tumors and provide a reliable, accurate, and independent prognostic and predictive marker of tumor biology in the post-genomic era (13).

The prognostic or predictive markers are usually associated, although not necessarily exclusively, with histologic examination of tumor tissue, whereas early detection biomarkers are measured in more readily accessible fluids, which come in contact with potential tumors. The value in making this distinction is to indicate the purpose for which the marker will be used. The same marker may in fact serve multiple purposes. For example, CA-125 can be used for prognosis (rate of decline of CA-125 after surgery is correlated with survival; ref. 14) and for early detection of recurrence (15), and it also is being investigated in prospective clinical trials as a marker for early detection of new cancers (4).

This review will focus on biomarkers of early detection and markers of early recurrence or relapse.

**APPLICATIONS OF RENAL CELL CARCINOMA BIOMARKERS FOLLOWING CLINICAL DIAGNOSIS OF RENAL CELL CARCINOMA**

Molecular renal cell carcinoma biomarkers, namely disease-specific changes in urine or plasma proteins or DNA profiles that correlate with local and/or distant disease relapse or progression, would be very useful guides for thera-
BIOLOGICAL MATERIALS FOR BIOMARKER INTERROGATION

Several human cell lines are available for study of clear cell renal cell carcinoma (16). Many of these cell lines display inactivation of von Hippel-Lindau (VHL) tumor suppressor gene. Inactivation of VHL is a hallmark in most sporadic clear cell renal tumors, and it occurs early in renal carcinogenesis. Reconstitution of VHL activity in these cell lines leads to suppression of their ability to grow as tumors in the mouse xenograft assay (17). Consequently, paired isogenic cell lines that lack or have reconstituted VHL can offer an in vitro comparison of the molecular events associated with early steps of VHL-related renal carcinogenesis.

Comparison of gene expression profiles of VHL+/+ with VHL−/− cell lines may identify secreted proteins that could serve as candidate molecular markers. Alternatively, the profile of secreted proteins could be directly interrogated by analyzing comparatively the secreted proteome of these cell lines in conditioned tissue culture supernatants. A combination of genomic and proteomic analysis appears reasonable because it is possible that not every difference in genomic messages will translate into differences at the protein level. The advantages of such an in vitro approach consists in the availability of these cell lines, the ease of their manipulation, and the relative low cost, at present, required for maintenance and manipulation. The pitfalls are obvious: information obtained from cell lines may harbor artifacts derived from additional mutations, not present in the original tumor, and favored by the tissue culture conditions or polymorphisms that correspond to the individual tumor characteristics from which the cell line was derived. Thus, analysis of several lines is warranted before compiling an initial list of candidate markers. In addition, cell line-based biomarker discovery does not reflect changes attributed to host-tumor interaction; the latter could produce tumor cell-independent molecular markers that reflect disease activity. Some of these shortcomings may be overridden by analyzing tumor specimens directly. Proteomic analysis of cell lines and renal tumor specimens is just beginning. Koteny et al. (18) provided an example of such an approach, and they identified nuclear matrix proteins specifically up-regulated in cell lines and human renal cancer. It is possible that the presence of these proteins can be captured quantitatively in serum or urine of renal cell carcinoma patients and that SEM proteins will serve as biomarkers for early detection of renal cell carcinoma.

Tumor development is the outcome of a series of defined genetic events. Among them is loss of tumor suppressor genes (indicated often by loss of heterozygosity in the genomic areas encompassing the gene) and mutational activation or amplification of proto-oncogenes. Comparative genomic hybridization and array analysis for single nucleotide polymorphisms between tumor samples and the corresponding normal tissue may reveal such critical events that characterize renal carcinogenesis (19, 20). This information, coupled with proteomic analysis of tumor samples compared with normal kidney parenchyma, may highlight proteins that can serve as candidate molecular renal cell carcinoma markers.

Genomic and proteomic analysis of tumor specimens reflects tumor-host interactions and, in addition, may provide stage-specific information. Great caution should be exerted, however, in processing of the tumors and their storage conditions because these may critically affect RNA and protein preservation (21). Standardized tissue banking is currently a severe limitation in the availability of this biological material, and its nationwide establishment constitutes a major effort coordinated by the NIH and the Early Detection Research Network. An example of a candidate marker for ovarian cancer identified by mRNA expression level analysis of tumor cell lysate compared with normal human ovarian surface epithelium lysate is prostasin (22).

An alternative approach, which complements cell line- and tumor-based discovery of molecular biomarkers, seeks to identify quantitative or qualitative changes of circulating molecules directly in plasma and/or urine of renal cell carcinoma patients.

A straightforward approach is to search for molecular markers present in plasma or urine of renal cell carcinoma patients before but not after nephrectomy for localized renal cell carcinoma. The technology for this approach is now available. Plasma is a complex biological fluid and, depending on the strategy of analysis (see below), may allow for a variable threshold for detection of changes in protein concentration (variable dynamic range).

Nevertheless, plasma or urine proteomics emerges as an attractive strategy for biomarker detection for several reasons. Plasma and urine are biological compartments that can be sampled easily and serially. Changes in plasma or urine protein profiles reflect not only the proteins secreted by the tumor cells but also tumor-host environment interactions. They also incorporate parameters such as stability and metabolism of secreted proteins that are viewed as candidate markers by in vitro approaches (proteomics of cell lines and tumor samples).

PROTEOMIC STRATEGIES FOR EARLY IDENTIFICATION OF BIOMARKERS

There are currently complementary strategies for profiling a complex protein sample derived from tumor cell lines, tumor specimen, or a biological fluid such as plasma or urine. These strategies are based on protein profile comparison between biological fluids obtained from cancer patients and control healthy adults or patients with nonmalignant conditions (23). Simple fluids such as urine can be analyzed by precipitation of a total protein content, followed by either resolution of these proteins by two-dimensional gel electrophoresis and identifica-
tion of tumor-associated proteins (24) or tandem mass spectrom-
etry of the whole protein content (25). For more complex fluids such as plasma or tumor specimens, there is likely a need to first fractionate the proteins by liquid chromatography based on size or other physicochemical properties such as isoelectric focusing. Each fraction can then be resolved by one- or two-dimensional protein gel electrophoresis. Comparison between profiles obtained in the presence or absence of the disease may reveal unique proteins or quantitative protein changes or lead to identification of protein modifications associated with the presence of the disease. Visualized proteins can be excised from the gel and identified by mass spectroscopy. Laser capture microdissection of tumor specimen may provide highly homogeneous cell samples to be compared with normal tissue parenchyma (26).

A recently developed technique is based on an initial identification of a protein pattern (without knowledge of the protein identity) that correlates with the presence and/or recurrence of the disease. In the matrix-assisted laser desorption/ionization time-of-flight technique, proteins are captured en mass on a solid surface with the help of an energy-absorbing matrix. Laser excitation or energy transfer results in fragmentation of the proteins. An anode-placed electrode detects ionized peptides and the time of flight to the electrode translates into calculated mass. Statistical and computer science techniques applied to the intensity levels measured at each mass-charge ratio may identify patterns that correlate with presence of the disease. A similar approach takes advantage of the possibility to create surfaces that selectively capture proteins that interact preferentially with hydrophobic, hydrophilic, ionic, metal binding, or mixed property surfaces and therefore enrich for a selected subtype of proteins for profile analysis surface-enhanced laser desorption ionization (SELDI).

The matrix-assisted laser desorption/ionization/SELDI-time-of-flight approach can be combined with fractionation, as described above, to identify proteins of interest. Matrix-assisted laser desorption/ionization-SELDI analysis can pinpoint the fractions of interest (fractions that contain novel protein patterns) in complex biological fluids and help analyze these fractions derived by protein gel electrophoresis. Alternatively, specific fractions of interest can be subjected to tandem mass spectrometric analysis that may provide sequence identification of the peptides. These techniques have displayed power in detecting disease-associated patterns and identifying candidate molecular markers (27–30).

A critical evaluation of proteomic techniques fleshed out key issues that affect the potential clinical application of these approaches. It appears that storage conditions, calibration of the instruments, and potential variability in the properties of protein-capturing surfaces may introduce significant variability in protein pattern formation, all providing challenges to be addressed in the future (21).

PROPOSED ALGORITHM FOR BIOMARKER DISCOVERY

Analysis of cell lines, tumor specimens, and biological fluids is only the first step in the identification of a biomarker. For example, such an analysis could conceivably provide three overlapping lists of proteins that appear to be deregulated in renal carcinoma cells, or tumor samples, or their concentration is changed significantly in plasma of renal patients after nephrectomy. Only a restricted number of proteins could be processed for further validation. A priority score/algorithm of approach needs to be created for the candidate biomarkers to identify which are deserving of additional investigation (Fig. 2). Candidate biomarkers with higher priority scores would likely be the ones present in all three lists, the ones that display a tissue restricted pattern of expression (kidney tissue), and presenting a wide range of change in the presence of the tumor. In addition, candidate markers for which high-quality antibodies are available would be of high priority as these are relatively easy to test.

The criteria proposed herein may vary, of course, depending on the nature of the protein or the biology of the disease. Once a short list is formed, antibodies and assays can be developed that recognize the candidate marker specifically and avidly in the appropriate biological compartment (e.g., antibodies against a plasma circulating protein should display minimal serum interference). Candidate biomarker changes, as detected initially in identification samples, should repeat themselves in test samples, collected independently.

The last stages will require the development of a statistical rule for combining biomarkers to achieve a maximum specificity and sensitivity and the prospective validation of this rule in an independent patient population.
OPEN DISCUSSION

**Dr. William Kaelin:** There have been questions about this SELDI technology in terms of how robust the signatures are and how sensitive they are to the methods by which the samples are procured and stored and what day they’re analyzed. Have you satisfied yourself that SELDI is appropriate technology for looking for markers?

**Dr. Othon Iliopoulos:** There are limitations to this technology. The first limitation is that the range of changes/dynamic range is more limited in the SELDI technology. The second is that SELDI yields just protein patterns associated with the presence of the disease, and you can either preliminarily validate your pattern and use it for diagnostic purposes as such or guess what your protein is and then do your fractionation analysis to identify and validate the molecular marker. SELDI technology has been very useful in producing patterns. To identify potential targets, I would use more targeted fractionation technology, and I would start from the most simple and go to the most complex compartments.

**Dr. Richard Childs:** Based on the limitations of SELDI-TOF protein pattern recognition as a screening test, even if you were to look at a high-risk population, do you think that this technology, at least the way that it exists today, could be used diagnostically?

**Dr. Iliopoulos:** I think this is discovery technology. It has been producing reproducible patterns in ovarian cancer and getting some results in lung cancer, so I definitely think that it is a viable technology for pattern reproduction. I don’t think it is diagnostic technology.

**Dr. Childs:** I think one of the problems with proteomic pattern recognition is that after you come up with the test set, you must go back to validate it against controls, making the test numbers extraordinarily large. Also, because of the different diseases that might mimic that process, it seems a difficult task in terms of just using proteomic pattern recognition.

**Dr. Iliopoulos:** When we go to this type of marker identification, study structure is important. It is very important to...
incorporate the appropriate controls if we want to ask the right question. At the Dana-Farber Harvard Cancer Center Renal Cancer Program, individuals are used as their own controls and compared with themselves pre- and postsurgery.

**Dr. Michael Atkins:** There are some data coming out in other tumors, indicating that circulating methylated DNA patterns may actually yield signatures for various diseases with specific methylation gene patterns. Is that something that could be looked at in renal cancer?

**Dr. Iliopoulos:** People have looked at this. I think it is a very viable and reasonable approach. It has potentially the same limitations as the protein approach. It may not be specific.

**Dr. Robert Flanigan:** What would be your advice, for example, if we were going to do a phase III trial in renal cell cancer looking at an agent versus control? Is there any strategy from a biomarker perspective that you would recommend?

**Dr. Iliopoulos:** I think that the biomarker story is still in the discovery era, so it is too preliminary to go into a phase III trial. What will be very useful for the prefinal step of validation is to have stored samples and to be able to use them as an independent collection for preliminary marker validation. One of the least available and probably most easily accessible resources is good urine collection.

**Dr. Atkins:** Do we know enough about the best way to collect urine so that it can be studied? Can that type of approach be standardized and distributed to cooperative groups or other groups involved in large-scale trials?

**Dr. Iliopoulos:** I recommend not storing samples at −20°C. Urine should be stored at either −80°C or in your liquid nitrogen. Every time you store them at −20°C, things decay. In addition, plasma is a more reliable compartment than serum because serum has clotted and can mask angiogenic conditions. So, if somebody collects blood, I suggest collecting plasma.

### REFERENCES

Molecular Markers for Early Detection of Renal Carcinoma: Investigative Approach

Steven Skates and Othon Iliopoulos

_Clin Cancer Res_ 2004;10:6296S-6301S.

Updated version  Access the most recent version of this article at: http://clincancerres.aacrjournals.org/content/10/18/6296S

Cited articles  This article cites 27 articles, 8 of which you can access for free at: http://clincancerres.aacrjournals.org/content/10/18/6296S.full#ref-list-1

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.