Published in final edited form as:

Lancet Oncol. 2011 February 1; 12(2): 137–143. doi:10.1016/S1470-2045(10)70296-5.

Development and prospective evaluation of a 20-gene model for molecular nodal staging of bladder cancer

Steven Christopher Smith, M.D.^{1,^,†}, Alexander Spyridon Baras, Ph.D.^{2,^}, Garrett Dancik, Ph.D.¹, Yuanbin Ru, Ph.D.¹, Kuan-Fu Ding, B.S.³, Christopher A Moskaluk, M.D.², Yves Fradet, M.D.⁴, Jan Lehmann, M.D.⁵, Michael Stöckle, M.D.⁶, Arndt Hartmann, M.D.⁷, Jae K Lee, Ph.D.³, and Dan Theodorescu, M.D., Ph.D.^{1,8,*}

¹Department of Surgery, University of Colorado, Aurora, Colorado, USA

²Department of Pathology, University of Virginia, Charlottesville, Virginia, USA

³Department of Public Health Sciences, University of Virginia, Charlottesville, Virginia, USA

⁴Department of Urology, Laval University, Québec, CANADA

⁵Urology Office Prüner Gang, Kiel, GERMANY

⁶Department of Urology, Saarland University, Homburg/Saar, GERMANY

© 2011 Elsevier Ltd. All rights reserved.

*Correspondence to: Dan Theodorescu, Professor of Surgery and Pharmacology, and Director, University of Colorado Comprehensive Cancer Center, 13001E, 17th Pl. MS #F-434, Aurora Colorado, 80045, dan.theodorescu@ucdenver.edu, Tel. +1 (303) 724-7135 Fax. +1 (303) 724-3162. The corresponding author had full access to all of the data and the final responsibility to submit for publication.

[†]Current address: Department of Pathology, 4211 Medical Science I, 1301 Catherine, SPC 5602, Ann Arbor, MI 48109

[^]Equal contribution

Professorship Information:

Dan Theodorescu is a full Professor of Surgery and Pharmacology and Director of the US National Cancer Institute (NCI)-designated Comprehensive Cancer Center of the University of Colorado, Aurora, CO, USA.

Christopher Moskaluk is a full Professor of Pathology and Biochemistry and Molecular Genetics at the University of Virginia, Charlottesville, VA USA.

Yves Fradet is a full Professor of Surgery at Laval University, Québec, QC, Canada.

Arndt Hartmann, MD is full Professor and Chairman, Department of Pathology, University of Erlangen Nuremberg, Germany.

Michael Stöckle, M.D. is Professor and Chairman, Department of Urology, Saarland University, Homburg, Germany

Jae K Lee PhD is Professor and Chief of the Division of Biostatistics, Department of Public Health Sciences, University of Virginia, Charlottesville, VA.

Authors' Contributions: Study designed by SCS, ASB, DT. Data were gathered and analyzed by SCS, ASB, GD, YR, KFD, CAM, YF, JL, MS, AH, JKL and DT and manuscript written by SCS, ASB, GD, YR and DT with contributions and final approval of all.

Conflict of Interests Statement: DT, JKL, SCS and ASB are listed as inventors on a provisional patent application based on the technology outlined in this work.

Ethics committee approval: Work was approved by the IRBs of centers providing tissues (Laval University, Québec, CANADA; Institute of Pathology, University of Erlangen, GERMANY).

Systematic Review: Evidence supporting the rationale for performing this study was garnered from searches of the PUBMED database, employing management of muscle invasive urothelial cancer, neoadjuvant chemotherapy, and practice patterns in Urologic Oncology as keywords (reviewed recently in Hussain *et al.* JCO 2009). Most importantly, we interpreted Level 1 evidence in favor of the use of neoadjuvant cisplatin-based chemotherapy as sound given its having been recently reviewed and supported in a meta-analysis of all relevant trials by the COCHRANE DATABASE (Neoadjuvant chemotherapy for invasive bladder cancer. Cochrane Database Syst Rev. 2005; (2): CD005246).

Interpretation: Despite the preponderance of evidence in its favor, neoadjuvant chemotherapy, which results in a small, significant increase in survival in bladder cancer, is used infrequently due to inability to risk stratify patients before definitive surgical staging. This study proves the principle that molecular staging before surgery can change the way we view urothelial cancer management and practice by assessing, *a priori*, staging parameters before surgery, after which only adjuvant chemotherapy (unsupported by any Level 1 evidence) remains an option.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

⁷Institute of Pathology, University of Erlangen, GERMANY

⁸University of Colorado Comprehensive Cancer Center, Aurora, Colorado, USA

Abstract

Background—Neoadjuvant chemotherapy prior to cystectomy confers a survival benefit in bladder cancer. Yet, it has not been widely adopted, since most patients do not benefit and we are currently unable to predict those that do. As the most important predictor of recurrence following cystectomy is pathologically positive nodes, we propose that tools defining this stage would be useful in selecting patients for neoadjuvant chemotherapy.

Methods—We developed a gene expression model (GEM) predictive of pathological node status for use on primary tumor tissue from clinically node negative (cN0) patients. From a subset of transcripts detected faithfully by microarrays from both paired frozen and formalin fixed tissues (N=32 pairs), we developed both the GEM and cutoffs identifying patient strata with elevated risk of nodal involvement using two separate training cohorts (N=90 and N=66). We then evaluated the GEM and cutoffs to predict node positive disease in tissues from a Phase III trial cohort (AUO-AB-05/95, N=185).

Findings—A 20-gene GEM was developed and exhibited an AUC=0.67 [95%CI 0.59–0.75], for prediction of nodal disease at cystectomy in AUO-AB-05/95. The cutoff system identified subjects with high (RR 1.74, [1.03–2.93]) and low (RR 0.70, [0.51–0.96]) relative risk of node positive disease. Multivariate logistic regression demonstrated the GEM predictor was independent of age, gender, pathologic stage, and lymphovascular space invasion (P=0.019).

Interpretations—Selecting patients for neoadjuvant chemotherapy based on risk of node positive disease has the potential to benefit high-risk patients while sparing others toxicity and delay to cystectomy.

Funding—U.S. National Cancer Institute (R01CA143971)

Keywords

bladder cancer; metastasis; staging; molecular diagnostics

INTRODUCTION

One third of bladder cancer patients present with muscle invasive disease (MIBC, stage T2) (1), which when clinically confined to the bladder (stage N0M0), are treated by radical cystectomy and pelvic lymphadenectomy or chemoradiation (2). Despite this, up to 50% of such patients develop metastatic disease, which is nearly always fatal (3). Fortunately, evidence from randomized control trials (4, 5) and a meta-analysis thereof (6) support use of neoadjuvant, platinum-based combination chemotherapy, in improving survival. Despite such evidence, neoadjuvant chemotherapy has not been widely adopted. In a recent series, encompassing approximately 60% of patients in the United States diagnosed with MIBC, <2% of patients received neoadjuvant chemotherapy (7). Recent reviews have identified barriers to its implementation, including concerns for delay of surgery and risk of disease progression (4).

A study of recurrence post-cystectomy in >9,000 patients (untreated by chemo or radiotherapy) reveals that ~80% of cases with pathologic node positive disease recurred, while only ~30% of extravesical, pathologic node negative or ~20% of organ confined cases recurred(8). Using the 10% disease-free survival benefit at 5 years from neoadjuvant chemotherapy for bladder cancer calculated in the clinical trial meta-analysis (6) only 2 out of 100 node negative, organ confined patients would benefit from neoadjuvant

chemotherapy. In contrast, for the patients with nodal positive disease (pN1-3), ~8 out of 100 patients could benefit from this approach. Hence, if pathologic nodal status could be predicted before surgery, one could identify a high risk patient stratum with a greater likelihood of benefit.

Here we develop and validate a gene expression model (GEM), evaluable on the primary tumor that predicts pathologic node positive (pN1-3) disease at cystectomy in clinically node negative patients. Use of this GEM would allow selection of those most likely to benefit most from neoadjuvant therapy while avoiding overtreatment and delay. By virtue of its applicability to formalin fixed tumor material, this approach may increase the use of neoadjuvant chemotherapy of bladder cancer in the oncologic community.

METHODS

Patients and Tissue Samples

Three independent cohorts of patients who underwent cystectomy and lymphadenectomy for bladder cancer were used to develop and validate the gene expression model (GEM) for nodal prediction (Table 1). All patients were clinically node negative (cN0) prior to cystectomy and had complete pathologic staging information. Importantly, no patients in any cohort were treated by systemic chemotherapy (adjuvant, neoadjuvant, or otherwise) was administered prior to harvesting of tissue used in this study. Two studies employed archival formalin-fixed paraffin embedded (FFPE) tissues of cystectomy specimens, the *Laval Cohort* (N=95) from l'Hôpital de l'Hotel-Dieu at Laval University, Québec, Canada and the *AUO Cohort* (N=188), which was from a Phase III, multicenter randomized control trial comparing two adjuvant chemotherapy regimens, AUO-AB 05/95, conducted by the German Arbeitsgemeinschaft Urologische Onkologie (AUO) (9). Of the 327 patients reported in the AUO cohort, archival tissues could only be retrieved on 200 of which 185 produced nucleic acid extracts of sufficient quality for microarray analysis. Distributions of clinicopathologic variables including survival did not differ significantly between the subset profiled (N=185) and the non-profiled group (N=142) (see Supplementary Methods and Results). The third cohort of MIBCs (pT2), where tissues were preserved by fresh freezing (FF), were obtained from supplementary data for a prior publication by Sanchez-Carbayo *et al.* (10) at the Memorial Sloan-Kettering Cancer Center, New York, the *MSKCC Cohort* (N=66). Additional gene expression datasets described in Supplementary Methods and Results were used for **1**) adaptation of the GEM to transurethral resection (TUR) specimens; **2**) adaptation of the GEM to the use on formalin fixed paraffin embedded (FFPE) tumor rather than fresh frozen (FF) material to permit widespread applicability of the test; **3**) evaluation of the ability of the GEM to predict overall survival.

Gene Expression Profiling and Data Processing

FFPE archival patient tissues from the Laval and AUO cohorts described above were reviewed by certified pathologists, who selected and harvested representative areas containing tumor >80% using a biopsy instrument (# 33-31A P/25, Miltex Inc.) that retrieves a 1.5mm × 3mm tissue core from the FFPE block. These samples were used for nucleic acid extraction, verification, amplification, and hybridization to HG-U133 plus 2.0 arrays, with analysis by RMA (11), as reported (12). Of the samples, 90/95 Laval and 185/188 AUO provided high quality hybridizations for use in GEM development.

Development of a Gene Expression Model (GEM) Predictive of Pathological Node Status

All development and statistical analysis of the gene expression model (GEM) was performed in the statistical programming language, Matlab (The Mathworks, Natick, MA). The Laval Cohort and MSKCC Cohort were used for training and the AUO Cohort was used

strictly for independent validation (details in Supplementary Methods and Results). Before model development, two “filters” were applied. The first removed candidates that did not meet FFPE to FF high fidelity criteria (details in Supplementary Methods and Results), to ensure applicability to both FF and FFPE samples. The identified candidates that had similar expression between transurethral resection (TUR) and the cystectomy, an important step to ensure applicability of the test to the pre-cystectomy TUR specimen, were used. This step involved the comparison of gene expression profiles between cystectomy and TUR acquired specimens and is described in detail in Supplementary Methods and Results).

To develop the model, using only gene expression probes that passed the above filters, we calculated a Wilcoxon two-sample test P-value for each probe’s association with pathologic node status (pN1-3 versus pN0) in the Laval Cohort (training). We then ranked the probes from lowest P-value to highest P-value. Our model selection proceeded as follows: 1) Starting with the top 5 most significant probes we used these probes to predict node status in the MSKCC Cohort (also training, optimization) using a WNN classification algorithm rooted in Bayesian decision theory and assuming a baseline prevalence (prior probability) of 23% nodal positive disease at cystectomy (3), see Supplemental Methods and Results. 2) We calculated the AUC for its prediction and recorded the value. We then repeated steps 1) and 2) for the top 6 most significant probes, the top 7 most significant probes, the top 8 most significant probes, and so on through the top 150 most significant probes (an arbitrary cutoff number). We found that of models including the top 5 through top 150 probes, the top 21 performed best on the MSKCC Cohort (AUC=0.72, [95%CI 0.58–0.85]). This was selected as the final model. Hierarchical clustering was used for visualization, while significance was assessed at an α of 0.05 in non-parametric T-tests or random permutation tests, as appropriate.

Development of a Cutoff System to Facilitate Clinical Interpretation of GEM Predictions

Because WNN prediction algorithm outputs a prediction score between 0 and 1, difficult to interpret alone, we used another resampling-based cross-validation strategy in the training cohorts (Laval and MSKCC Cohorts), to prospectively develop thresholds based on relative risk (RR) of nodal positive disease to identify patient groups at statistically significantly higher or lower risk. These thresholds were defined based on the following criteria *a priori*: (1) with 95% confidence the relative risk in predicted high and low risk groups must be increased or decreased; (2) less than 15% of calls in a hypothetical population could be deemed intermediate; and (3) with 95% confidence, the relative risk of nodal positive disease in the intermediate prediction group must include unity (1). Of the thresholds that met our criteria, optimal high and low risk cutoffs were selected, equally weighting the relative risk for the positive and negative node prediction strata. A detailed description of these methods is included in Supplementary Methods and Results.

Independent Testing of the GEM and Relative Risk Cutoffs

After development of the GEM and RR-based cutoffs, both were tested for performance in the independent, AUO Cohort. Need to add here briefly. Additionally, multivariate logistic regression was used to test the GEM prediction scores for independent association with nodal status, adjusting for traditional clinical and pathologic tumor factors. To test for the association between the GEM genes and survival, a Cox Proportional Hazards Regression model was trained using survival data for both the Laval and MSKCC training cohorts in the statistical suite, R. This model was then tested for its ability to stratify the disease specific outcomes of the AUO cohort and another recently published cohort of muscle invasive bladder cancers (N=62) of from Chungbuk National University Hospital, South Korea, which was profiled on a separate microarray platform ((13)). For full details see Supplementary Methods and Results.

Statistical Methods

Microarray data extraction with RMA (11) and bootstrapping “high fidelity” analyses were performed in the Statistical Suite R, version 2.8.1 (www.r-project.org). Wilcoxon two-sample tests, model development, WNN prediction algorithm, receiver operator characteristic (AUC) testing, cutoff generation, cutoff testing, logistic regression analyses were performed in Matlab R2009B (The Mathworks, Natick, MA). AUC curves were plotted and 95% confidence intervals were calculated in Prism 5.0c (GraphPad Software, La Jolla, CA). Cox proportional hazards regression models were also built and tested in R.

RESULTS

Identification of transcripts reliably detected by microarray, independent of tissue preservation

Tissue preparation remains a significant challenge in the development of molecular diagnostics, particularly if the former are to be based on RNA species (14–16). To broaden the utility of an eventual clinical test based on work presented here, we identified RNA transcripts detected with high-fidelity independently of sample preservation (fresh freezing (FF) or formalin fixation with paraffin embedding (FFPE)), before development of our GEM. To that end, we used 32 paired FF and FFPE tissues profiled by microarrays to identify probes which were further refined to ensure they are expressed in TUR-acquired tumors as well as cystectomies (both steps described in Supplementary Methods and Results). These were used for GEM development on the Laval and MSKCC cohorts (Figure 1A).

Development of a GEM and cutoff system for predicting relative risk of pathologic node status in MIBC

Approximately 20–40% of patients with clinically organ confined MIBC have node positive disease at cystectomy(3). Per our **Methods**, we used microarray data from two (Laval and MSKCC) cohorts of cystectomy patients to develop a GEM predictive of nodal status, demonstrated in Figure 1B and detailed in Supplementary Methods and Results. We identified an optimal GEM consisting of probes for the top 20 genes (21 microarray probes) most significantly associated with node status in the Laval Cohort, based on their performance predicting the nodal status of the MSKCC Cohort (AUC=0.72, [95%CI 0.58–0.85]), Figure 1B). Figure 1C shows the 21 probes and their univariate association with pathological nodal status.

To provide clinical utility, we developed high and low-risk cutoffs of these continuous scores as has been done in other molecular tests (17). These were based on *relative risk* of nodal positive disease at cystectomy, and developed using a cross-validation approach in both training cohorts. Cutoffs were designed such that patients would be identified by the GEM as high-risk, low-risk, or intermediate risk, see **Methods** and Supplementary Methods and Results. Optimal cutoffs were chosen based on training data only, and yielded a low risk group exhibiting a relative risk of 0.47 [95%CI 0.31–0.70] and a high-risk group exhibiting a relative risk of 2.25 [1.18–4.28], Figure 2A.

An advantage of using both a FFPE (Laval) and FF-preserved (MSKCC) cohort for training is that this approach allows examination of the extent to which using the FFPE to FF high-fidelity transcripts facilitated the generalization of the model across the two differently preserved study groups. We evaluated the contribution of this technique to the efficacy of intercohort FFPE to FF prediction by examining the performance of the top 5 through 150 probes most associated with nodal status in the Laval Cohort (FFPE) with and without pre-selection of only high-fidelity probesets. We observed that high-fidelity probesets were

superior at prediction of nodal status in the MSKCC Cohort from the Laval Cohort, demonstrating this pre-selection methodology contributes to better intercohort prediction ($P<0.0001$) (Figure 2B).

Validation of the GEM and cutoff system on prospectively collected cohort

Ideally, to determine efficacy and avoid bias predictive biomarkers should be evaluated on cohorts collected prospectively. We accomplished this by testing the performance of the 20-gene model on tissues from patients enrolled in the AUO-AB 05/95 trial (AUO Cohort, $N=185$) (9). We observed a significant ability of the predictor to discriminate between pathologically nodal positive and nodal negative tumors ($AUC=0.67$ [95%CI 0.60 to 0.75], $P<0.0001$, Figure 3A). We also found the risk stratification scheme we developed generalized to this independent cohort; wherein high and low-risk groups with increased (RR 1.74 [1.03–2.93]) or decreased (RR 0.70 [0.51–0.96]) relative risk of nodal positive disease were identified, Figure 3B.

To have clinical value, molecular prediction strategies should show significantly improved prediction compared to standard clinicopathologic parameters. Starting from the Laval and MSKCC training cohorts, we found that of gender ($P=0.67$), age ($P=0.70$), and pathologic stage ($P=0.025$), only pathologic tumor stage was significantly associated with nodal status at cystectomy. Using univariate and bivariate logistic regression models built only using the Laval and MSKCC training cohorts, we compared the node status prediction performance of pathologic tumor stage alone *versus* pathologic tumor stage plus GEM prediction in the independent AUO cohort. We observed that pathologic tumor stage exhibited minimal prediction performance in relevant cases of the AUO cohort ($AUC=0.52$, 95%CI 0.42–0.61), whereas the combination of pathologic tumor stage plus GEM prediction score improved the combined prediction accuracy substantially ($AUC=0.65$, 95%CI 0.56–0.74), see Supplementary Methods and Results for full details. Strikingly, this incremental increase in the AUC persisted despite the use of pathologic staging parameters from cystectomy specimens, much more accurate and precise staging than the clinical stage determined from transurethral resection specimens, which may understage tumors in excess of 50% of cases (18), which would be the relevant comparison in practice.

We also tested the GEM prediction for independence of association from other clinicopathologic variables. Using multivariate logistic regression including age, gender, pathologic stage, lymphovascular space invasion (LVSI) at cystectomy, and the GEM score, we found that in the AUO test cohort the GEM score was significantly independently associated with nodal status $P=0.019$ (Table 2). Independence of the GEM beyond these variables is a very stringent test, as pathologic stage would not be available at the time GEM evaluation as would be applied in clinical practice, nor would definitive LVSI, which also may differ substantially between transurethral resection and cystectomy specimens (19).

Relationship of the GEM Model Genes to Patient Survival

Given the strong relationship of pathologic node status to disease free survival observed in several large series (3, 8), the corollary question arises as to whether the 20 node-related GEM genes could also model patient survival post-cystectomy. Such an observation would further support the relevance of these genes by relating them to another key prognostic parameter. We used multivariate Cox proportional hazards regression on the 20 GEM genes to model survival, again trained on the Laval and MSKCC training cohorts (see Supplementary Methods and Results). We found that the model genes could stratify survival in the AUO Cohort ($P=0.048$), as well as another cohort recently profiled by a separate gene expression profiling technology ($P=0.027$), supportive of the generality of their association with poor prognosis.

DISCUSSION

We developed a test that can predict a powerful determinant of prognosis following cystectomy, namely node positive disease (3, 8), with the hope that such molecular intelligence, for which no other molecular marker exists, would provide a tool allowing more effective and frequent use of neoadjuvant chemotherapy by selection of high-risk patients and avoidance of overtreatment of low-risk patients. Our use of multiple cohorts sourced from different institutions on different continents, including a prospective, multicenter cohort from a randomized control trial supports the generalizability of the approach across very different populations (20). This constitutes, to our knowledge, the largest genome-wide profiling study of bladder cancer to date (275 patients), and first report profiling or generating a biomarker from FFPE bladder tissues. Last, we hope that identification of high-fidelity transcripts preserved from tissues processed either by FFPE or FF will be of use to other research groups in development of gene expression-based molecular biomarkers on FFPE material given its widespread use.

Most importantly, this tool demonstrated the ability to significantly predict node positive and node negative patients when tested on prospectively-collected tissues from the AUO cohort, with a level of performance that is similar to that of gene expression based molecular predictors currently in clinical use for other tumor types(21). Notably, this evaluation was independent of classic clinical and pathologic predictive factors, the latter being the current “gold standard” variables in determination of prognosis after cystectomy. To facilitate use of this tool in clinical practice, we developed, based on training data alone, a system of two cutoffs in the predictor’s scores, designed to identify groups with significantly increased or decreased risk of nodal involvement, relative to a baseline rate of ~25% node positive disease, taken from a published series of >1000 patients, (3). These cutoffs again performed as intended when applied to the independent AUO Cohort. The reported AUC of 0.67 allowed us to develop a system of cutoffs, identifying clinically actionable populations, with potential clinical utility far beyond the raw overall performance curve

These results come with caveats. Our GEM was evaluated on samples obtained at cystectomy rather than transurethral resection (TUR), the latter being the technique used to obtain diagnostic tissue before neoadjuvant chemotherapy and upon which the GEM would be applied in clinical practice. While studies have successfully validated on TUR specimens GEMs developed from cohorts containing both cystectomies and TUR specimens (22) or TUR specimens alone (23), we went further to ensure that our GEM was suitable for use on TUR specimens. First, we ensured that the genes from which the model was built had similar expression patterns in both TUR and cystectomy specimens. Second, the tissue used here for gene expression sampling was harvested with a biopsy instrument that retrieves a 1.5 mm tissue core from an FFPE block containing part of the bladder tumor. Arguably, such minute sampling is equally prone to intratumoral variability of gene expression whether it is done from a cystectomy specimen or the diagnostic TUR, since in both cases the core sample that would be used for profiling is very small compared to the overall tumor bulk.

Further, the optimal extent of intraoperative lymphadenectomy remains controversial for bladder cancer(24). Given the large numbers of patients in this study and their international distribution, it is likely that differences in the node dissection were present in the training sets compared to the validation set. However, despite this it is remarkable that when the model was assessed against AUO cohort patients enrolled in a Phase III explicitly mandating a complete node dissection for enrollment (9), the GEM exhibits significant performance, independent of classic clinical and pathologic variables, and similar in performance to molecular predictors used for other tumor types. In addition, patients found to have

pathologically negative nodes despite a strong GEM prediction to the contrary, may indeed harbor metastasis not detected by conventional pathological node assessment. We are considering testing this issue directly in future studies, in cases where archival nodal tissues remain by using sensitive molecular methods for cancer cell detection (25) in pathologically negative nodes of tumors called high risk by our GEM assay.

As newer imaging modalities (26, 27) find a role in clinical practice, future studies should compare their performance to that of the GEM described here. Indeed, future multivariate models could be built and validated using GEMs, novel imaging technologies coupled with pathological findings on the transurethral resection specimen (such as LVSI) to provide superior performance to any individual technique. We also see the opportunity to use data obtained from microarray profiling of tumor tissue not only for prediction of node status but for prediction of response to chemotherapy and selection of most appropriate drugs (23, 28–30). In summary, this nodal prediction model constitutes an important clinical advance in the personalization of bladder cancer therapy by selecting those at high risk for recurrence and thus likely best candidates for neoadjuvant chemotherapy. Currently, such information is lacking, leading to a gap between evidence and practice (24). It is our hope that assays like the one described herein could reduce this gap.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Role of Funding Source: The study was funded by the National Cancer Institute (NCI) of the US National Institutes of Health (R01CA143971). The NCI had no role in design, analysis, interpretation, or writing of the manuscript.

References

1. Botteman MF, Pashos CL, Redaelli A, Laskin B, Hauser R. The health economics of bladder cancer: a comprehensive review of the published literature. *Pharmacoeconomics*. 2003; 21(18):1315–30. [PubMed: 14750899]
2. Kaufman DS, Shipley WU, Griffin PP, Heney NM, Althausen AF, Efird JT. Selective bladder preservation by combination treatment of invasive bladder cancer. *N Engl J Med*. 1993 Nov 4; 329(19):1377–82. [PubMed: 8413433]
3. Stein JP, Lieskovsky G, Cote R, Groshen S, Feng A-C, Boyd S, et al. Radical Cystectomy in the Treatment of Invasive Bladder Cancer: Long-Term Results in 1,054 Patients. *J Clin Oncol*. Feb 1; 2001 19(3):666–75. [PubMed: 11157016]
4. Grossman HB, Natale RB, Tangen CM, Speights VO, Vogelzang NJ, Trump DL, et al. Neoadjuvant chemotherapy plus cystectomy compared with cystectomy alone for locally advanced bladder cancer. *N Engl J Med*. 2003 Aug 28; 349(9):859–66. [PubMed: 12944571]
5. Hall RR. Updated results of a randomized trial of neoadjuvant cisplatin (C) methotrexate (M) and vinblastine (V) chemotherapy for muscle-invasive bladder cancer. *Proc Am Soc Clin Oncol*. 2002; 21:178a, Abstract 710.
6. Neoadjuvant chemotherapy for invasive bladder cancer. *Cochrane Database Syst Rev*. 2005; (2):CD005246. [PubMed: 15846746]
7. David KA, Milowsky MI, Ritchey J, Carroll PR, Nanus DM. Low incidence of perioperative chemotherapy for stage III bladder cancer 1998 to 2003: a report from the National Cancer Data Base. *J Urol*. 2007 Aug; 178(2):451–4. [PubMed: 17561135]
8. Bochner BH, Kattan MW, Vora KC. Postoperative nomogram predicting risk of recurrence after radical cystectomy for bladder cancer. *J Clin Oncol*. 2006 Aug 20; 24(24):3967–72. [PubMed: 16864855]

9. Lehmann J, Retz M, Wiemers C, Beck J, Thuroff J, Weining C, et al. Adjuvant cisplatin plus methotrexate versus methotrexate, vinblastine, epirubicin, and cisplatin in locally advanced bladder cancer: results of a randomized, multicenter, phase III trial (AUO-AB 05/95). *J Clin Oncol*. 2005 Aug 1; 23(22):4963–74. [PubMed: 15939920]
10. Sanchez-Carbajo M, Socci ND, Lozano J, Saint F, Cordon-Cardo C. Defining molecular profiles of poor outcome in patients with invasive bladder cancer using oligonucleotide microarrays. *J Clin Oncol*. 2006 Feb 10; 24(5):778–89. [PubMed: 16432078]
11. Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP. Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res*. 2003 Feb 15; 31(4):e15. [PubMed: 12582260]
12. Bosman FT, Yan P, Tejpar S, Fiocca R, Van Cutsem E, Kennedy RD, et al. Tissue biomarker development in a multicentre trial context: a feasibility study on the PETACC3 stage II and III colon cancer adjuvant treatment trial. *Clin Cancer Res*. 2009 Sep 1; 15(17):5528–33. [PubMed: 19690194]
13. Kim WJ, Kim EJ, Kim SK, Kim YJ, Ha YS, Jeong P, et al. Predictive value of progression-related gene classifier in primary non-muscle invasive bladder cancer. *Mol Cancer*. Jan 8; 9(1):3. [PubMed: 20059769]
14. Hewitt SM, Lewis FA, Cao Y, Conrad RC, Cronin M, Danenberg KD, et al. Tissue handling and specimen preparation in surgical pathology: issues concerning the recovery of nucleic acids from formalin-fixed, paraffin-embedded tissue. *Arch Pathol Lab Med*. 2008 Dec; 132(12):1929–35. [PubMed: 19061293]
15. Toussaint J, Sieuwerts AM, Haibe-Kains B, Desmedt C, Rouas G, Harris AL, et al. Improvement of the clinical applicability of the Genomic Grade Index through a qRT-PCR test performed on frozen and formalin-fixed paraffin-embedded tissues. *BMC Genomics*. 2009; 10:424. [PubMed: 19744330]
16. van Laar RK, Ma XJ, de Jong D, Wehkamp D, Floore AN, Warmoes MO, et al. Implementation of a novel microarray-based diagnostic test for cancer of unknown primary. *Int J Cancer*. 2009 Sep 15; 125(6):1390–7. [PubMed: 19536816]
17. Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, et al. A Multigene Assay to Predict Recurrence of Tamoxifen-Treated, Node-Negative Breast Cancer. *N Engl J Med*. Dec 30; 2004 351(27):2817–26. [PubMed: 15591335]
18. Cheng L, Neumann RM, Weaver AL, Cheville JC, Leibovich BC, Ramnani DM, et al. Grading and staging of bladder carcinoma in transurethral resection specimens. Correlation with 105 matched cystectomy specimens. *Am J Clin Pathol*. 2000 Feb; 113(2):275–9. [PubMed: 10664630]
19. Kunju LP, You L, Zhang Y, Daignault S, Montie JE, Lee CT. Lymphovascular invasion of urothelial cancer in matched transurethral bladder tumor resection and radical cystectomy specimens. *J Urol*. 2008 Nov; 180(5):1928–32. discussion 32. [PubMed: 18801529]
20. Shariat SF, Lotan Y, Vickers A, Karakiewicz PI, Schmitz-Drager BJ, Goebell PJ, et al. Statistical consideration for clinical biomarker research in bladder cancer. *Urol Oncol*. Jul-Aug; 28(4):389–400. [PubMed: 20610277]
21. Marchionni L, Wilson RF, Wolff AC, Marinopoulos S, Parmigiani G, Bass EB, et al. Systematic Review: Gene Expression Profiling Assays in Early-Stage Breast Cancer. *Annals of Internal Medicine*. Mar 4; 2008 148(5):358–69. [PubMed: 18252678]
22. Lindgren D, Frigyesi A, Gudjonsson S, Sjodahl G, Hallden C, Chebil G, et al. Combined gene expression and genomic profiling define two intrinsic molecular subtypes of urothelial carcinoma and gene signatures for molecular grading and outcome. *Cancer Res*. May 1; 70(9):3463–72. [PubMed: 20406976]
23. Als AB, Dyrskjot L, von der Maase H, Koed K, Mansilla F, Toldbod HE, et al. Emmprin and survivin predict response and survival following cisplatin-containing chemotherapy in patients with advanced bladder cancer. *Clin Cancer Res*. 2007 Aug 1; 13(15 Pt 1):4407–14. [PubMed: 17671123]
24. Hussain MH, Wood DP, Bajorin DF, Bochner BH, Dreicer R, Lamm DL, et al. Bladder cancer: narrowing the gap between evidence and practice. *J Clin Oncol*. 2009 Dec 1; 27(34):5680–4. [PubMed: 19858384]

25. Kurahashi T, Hara I, Oka N, Kamidono S, Eto H, Miyake H. Detection of micrometastases in pelvic lymph nodes in patients undergoing radical cystectomy for locally invasive bladder cancer by real-time reverse transcriptase-PCR for cytokeratin 19 and uroplakin II. *Clin Cancer Res.* 2005 May 15; 11(10):3773–7. [PubMed: 15897575]
26. Kibel AS, Dehdashti F, Katz MD, Klim AP, Grubb RL, Humphrey PA, et al. Prospective study of [18F]fluorodeoxyglucose positron emission tomography/computed tomography for staging of muscle-invasive bladder carcinoma. *J Clin Oncol.* 2009 Sep 10; 27(26):4314–20. [PubMed: 19652070]
27. Thoeny HC, Triantafyllou M, Birkhaeuser FD, Froehlich JM, Tshering DW, Binser T, et al. Combined ultrasmall superparamagnetic particles of iron oxide-enhanced and diffusion-weighted magnetic resonance imaging reliably detect pelvic lymph node metastases in normal-sized nodes of bladder and prostate cancer patients. *Eur Urol.* 2009 Apr; 55(4):761–9. [PubMed: 19144456]
28. Lee JK, Havaleshko DM, Cho H, Weinstein JN, Kaldjian EP, Karpovich J, et al. A strategy for predicting the chemosensitivity of human cancers and its application to drug discovery. *Proc Natl Acad Sci U S A.* 2007 Aug 7; 104(32):13086–91. [PubMed: 17666531]
29. Smith SC, Baras AS, Lee JK, Theodorescu D. The COXEN principle: translating signatures of in vitro chemosensitivity into tools for clinical outcome prediction and drug discovery in cancer. *Cancer Res.* Mar 1; 70(5):1753–8. [PubMed: 20160033]
30. Williams PD, Cheon S, Havaleshko DM, Jeong H, Cheng F, Theodorescu D, et al. Concordant gene expression signatures predict clinical outcomes of cancer patients undergoing systemic therapy. *Cancer Res.* 2009 Nov 1; 69(21):8302–9. [PubMed: 19843853]

Figure 1. Development of a GEM predictor of pathological nodal status at cystectomy

A. High fidelity transcript discovery: For ease of implementation, we first used Affymetrix HG-U133 plus 2.0 microarray data for a cohort of 32 paired tissues that had been preserved by formalin fixation, paraffin embedding (FFPE) and fresh freezing (FF) to develop a set of probesets detected with high-fidelity by either means of tissue preservation. After 1000-fold bootstrapping the correlation of probes across the paired tissues, we selected probes maintaining positive correlation at the 2.5th percentile. 12,402 of the high-fidelity probes were common to the U133 plus 2.0 platform and the U133A platforms, and after ensuring these genes were not expressed differently in TUR specimens, these were used for GEM development. **B. Model development using training cohorts:** The GEM was developed by linear forward searching of probes most significantly correlated with nodal status in the Laval Cohort and predicting nodal status on the MSKCC Cohort, using a weighted nearest neighbor (WNN) predictor based on correlation of samples. The maximal AUC for discrimination was attained at 21 probes, representing 20 transcripts, (AUC=0.72, [95% CI 0.58–0.85]), and this model was selected as the final model based on training data alone. **C. The 20-gene model:** Univariate differences for each probe's expression across

nodal status are shown. The normalized expression in node positive samples (pN1-3) in the Laval Cohort is plotted in terms of standard deviations of difference compared to the node negative (pN0) samples of the Laval Cohort (error bars represent 95% confidence intervals).

Figure 2. A. Risk cutoff development for 20-gene model using training cohorts

Using a cross-validation approach applied to both Laval and MSKCC Cohorts, cutoffs were developed from the training cohorts to identify patient risk strata with significantly high or low relative risk of nodal positive disease. **B. Comparison of high fidelity vs. all**

transcript use in GEM based prediction of node positive disease: Box-whisker (boxes, median and interquartile range; whiskers 5th and 95th percentiles) plot of distributions of Area Under the ROC Curve performance of models based on the top 5 through 150 FFPE to FF high-fidelity node-associated genes compared to models based on the top 5 through 150 of all genes derived from the Laval cohort (FFPE). Predictions were made from the Laval Cohort (FFPE) on the MSKCC Cohort (FF). The distributions show a highly significant trend of superiority of the models based on high-fidelity genes, supporting the usefulness of this methodology. The top performing models are also plotted (solid triangles and AUCs for these shown).

Figure 3. Performance of the 20-gene model on a prospectively collected independent test cohort
A. Receiver Operating Characteristic analysis of the predictions outputted by the 20-gene model showed highly significant ability to discriminate between nodal negative and positive tumors in the AUO Cohort. **B.** Testing of the performance of the relative risk cutoffs on the independent test cohort found that the cutoffs developed from training data again identified groups within the AUO Cohort with significantly increased or decreased risk of nodal positive disease.

Table 1

Clinical and pathological characteristics of the patient cohorts used in the study

	Laval Cohort [±] N=90		MSKCC Cohort [*] N=66		AUO Cohort [^] N=185	
Age	Median	64	68.5		60	
	Range	46–81	37–85		34–71	
Gender		N	%	N	%	
	M	76	(84.4)	47	(71.2)	145 (78.4)
	F	14	(15.6)	19	(28.8)	37 (20.0)
	–	–	–	–	–	3 (1.6)
Grade	2	9	(10.0)	5	(7.6)	21 (11.4)
	3	81	(90.0)	61	(92.4)	163 (88.1)
	–	–	–	–	–	1 (0.5)
Stage	pT1	1	(1.1)	–	–	5 (2.7)
	pT2	12	(13.3)	10	(15.2)	25 (13.5)
	pT3	48	(53.3)	45	(68.2)	120 (64.9)
	pT4	29	(32.2)	11	(16.7)	31 (16.8)
	–	–	–	–	–	5 (2.7)
Node	pN0	17	(18.9)	41	(37.9)	82 (44.3)
	pN1	73	(81.1)	25	(62.1)	103 (55.7)

[±] The Laval Cohort referenced herein refers to an unpublished cohort of 90, archival specimens of which we profiled on the HG-U133 plus 2.0 microarray platform.

^{*} The MSKCC Cohort referenced herein refers to a subset of 66 muscle invasive cancer samples from a recent expression profiling study by Sanchez-Carbayo *et al.* (10), where frozen tumor samples were profiled on the HG-U133A microarray platform.

[^] The AUO Cohort referenced herein refers to 185 samples from a prior randomized control trial (AUO-AB-05/95) reported by Lehmann *et al.* (9), archival samples of which we profiled on the HG-U133 plus 2.0 microarray platform.

Table 2

Multivariate Logistic Regression of Clinicopathologic Parameters of GEM in the Test Cohort

Variable	Coefficient [±]	SE [±]	P-value [±]
Age [*]	0.0079	0.025	0.75
Gender	0.19	0.22	0.39
Stage [^]	0.13	0.44	0.77
LVSI [~]	0.80	0.36	0.030
GEM [#]	9.81	4.14	0.019

[±] Coefficient, standard error (SE) and P-value for independent association of indicated parameter with node status in included AUO test cohort cases, see Supplementary Methods and Results.

^{*} Age at cystectomy

[^] Pathologic stage at cystectomy (pT). TURBT clinical staging data are not available for this cohort.

[~] Pathologic lymphovascular space invasion, at cystectomy.

[#] Prediction score outputted by the GEM