Cystoscopy Revisited as the Gold Standard for Detecting Bladder Cancer Recurrence: Diagnostic Review Bias in the Randomized, Prospective CEFUB Trial

Madelon N. M. van der Aa, Ewout W. Steyerberg, Chris Bangma, Bas W. G. van Rhijn, Ellen C. Zwarthoff and Theo H. van der Kwast*

From the Departments of Pathology (MNMvdA, ECZ, THvdK), Urology (MNMvdA, CB, BWGvR) and Public Health (EWS), Erasmus Medical Center Rotterdam, Rotterdam, The Netherlands, and Department of Pathology, University Health Network (THvdK), Toronto, Ontario, Canada

Abbreviations and Acronyms

MA = microsatellite analysis

NMI = nonmuscle invasive

UC = urothelial carcinoma

Submitted for publication May 13, 2009.

Study received approval from participating institution ethical committees.

Supported by The Netherlands Organization for Health Research and Development Grant 945-02-046.

* Correspondence: Department of Pathology, University Health Network, 200 Elizabeth St., Toronto, Ontario M5G 2C4, Canada (telephone: 1 416 340 4597; FAX: 1 416 340 5517; e-mail: theo.vdkwast@uhn.on.ca).

For another article on a related topic see page 339.

Purpose: We evaluated the influence of knowledge of urine test outcome on the accuracy of cystoscopy (diagnostic review bias) during surveillance in patients with low grade, nonmuscle invasive urothelial carcinoma.

Materials and Methods: We performed a prospective, single-blind, randomized, multicenter clinical trial of surveillance by microsatellite analysis urine test in 448 patients with nonmuscle invasive (pTa, pT1, G1, G2) urothelial carcinoma. Positive or negative urine test results were only communicated to the urologist in the intervention arm of 226 patients, in which cystoscopy was done if the test was positive, and at 3, 12 and 24 months. Urine test results were not communicated in the control arm of 222 patients who underwent standard 3-month cystoscopy. The primary outcome measure was the number of histologically proven bladder cancer recurrences.

Results: At a median 34-month followup 218 recurrences were detected in the intervention arm compared to 163 in the control arm ($p < 0.001$). Of 131 cystoscopies done with knowledge of a positive urine test 42 recurrences were detected. Only 6 recurrences were found in the 120 cystoscopies done without information on the positive test result (chi-square $p < 0.001$). There was no difference in recurrence detection when urine test results were negative in the intervention and control arms (18 of 260 patients or 7% and 18 of 326 or 6%, respectively, $p = 0.45$).

Conclusions: Diagnostic review bias should be considered in the evaluation of point of care urine tests for bladder cancer monitoring. Awareness of a positive urine test result significantly improves the urothelial carcinoma detection rate using cystoscopy.

Key Words: urinary bladder, urothelium, carcinoma, bias (epidemiology), cystoscopy

LOOKING inside the hollow organs and body cavities of humans is a concept that was first put forward in 1806 by Bozzini. In 1877 Nitze (1848 to 1906) performed the first cystoscopy in Vienna.[1] Since then, cystoscopy has developed as the mainstay for diagnosing bladder disease with high sensitivity and specificity to detect papillary lesions.[2–4] The availability of flexible cystoscopes further improved its acceptance as a diagnostic tool. Nevertheless, alternative noninvasive diagnostic methods are being investigated to decrease cystoscopy frequency because cystoscopy is considered an in-

0022-5347/10/1831-0076/0
THE JOURNAL OF UROLOGY®

Vol. 183, 76-80, January 2010
Printed in U.S.A.
DOI:10.1016/j.juro.2009.08.150

vasive, time-consuming, costly procedure that is burdensome to the patient.[5,6] Implementation of urine tests such as MA to detect tumor cells in voided urine samples may represent an acceptable alternative to diagnose recurrent bladder UC and detect upper urinary tract UC.[7–9]

To detect flat lesions such as hyperplasia, dysplasia and carcinoma in situ white light cystoscopy has lower sensitivity than for papillary lesions. Nonetheless, the sensitivity of cystoscopy to detect papillary lesions may appear high since alternative diagnostic tests are not in routine use to identify UC recurrence. Thus, its alleged high sensitivity cannot be verified independently.

We performed a randomized, multicenter study of the safety and cost-effectiveness of surveillance in patients with low grade (1973 WHO grades 1 and 2) NMI UC using MA in voided urine samples as an alternative to cystoscopy.[10] In the control arm the urine test outcome was not communicated to the attending urologist at followup visits but in patients randomized to the test arm the urologist in charge was aware of the urine test outcome. We investigated whether diagnostic review bias may have caused the observation that the number of recurrences was higher in the test arm than in the control arm.

MATERIALS AND METHODS

Study Design

A total of 484 patients with primary or recurrent NMI UC (ie TNM 1997 stage pTa or pT1 and WHO 1973 grade 1 or 2) were recruited at 10 Dutch hospitals to participate in the randomized Cost-Effectiveness of Followup of Bladder Cancer Trial. Three of the 10 participating hospitals, including 1 community and 2 academic centers, have a urologist residency training program. A mean of 45 patients (range 8 to 128) was contributed by each hospital. White light cystoscopy was used for surveillance at the 10 hospitals. All participating urologists were informed about the literature on MA to detect recurrent UC.[6,7] Patients with a history of carcinoma in situ or grade 3 UC were excluded from participation. After transurethral resection of bladder tumor 448 patients were randomized after signing an informed consent form. Patients were stratified by hospital (10 hospitals), histopathological diagnosis (grade and stage) and *FGFR3* gene mutation status (mutation or WT). Participants were assigned to 2 trial arms by block randomization in order of appearance. Attending urologists received case record forms with a unique trial number linked to a database at the coordinating center.

The study began at the first cystoscopy 3 months after the transurethral resection of inclusion with followup by 3-month cystoscopy only (control arm) or by cystoscopy at 3, 12 and 24 months, and 3-month MA in voided urine samples (intervention arm). The MA urine test in voided urine samples was performed in each randomized group, as previously described in detail.[6,10] Laboratory analysis was done while blinded to clinical and histological status. Urine test results were only communicated by mail to the attending urologist in the test arm and according to protocol a positive test was followed by cystoscopy. The protocol did not prescribe random biopsies in the test arm when the urine test outcome was positive. In the test arm when a positive urine test outcome was incongruent with cystoscopy, at 2 subsequent followup visits the protocol recommended upper urinary tract imaging. The primary study outcome was tumor recurrence, defined as histopathologically confirmed UC detected by cystoscopy at followup. The ethical committee at participating institutions approved the study.

Statistical Analysis

Data analysis of randomized groups was performed on an intent to treat basis with the log rank test for Kaplan-Meier curves to compare time to recurrence. Patients with recurrence during followup remained in the same trial arm with a followup scheme similar to that after study inclusion. Subsequent recurrences were analyzed with modulated renewal.[11] Recurrence rates per followup visit were compared using the chi-square test. The randomized study was originally powered to show equivalence between randomized groups since followup in patients with MA in voided urine would be preferred if there were similar recurrence rates in the 2 arms. Equivalence was defined as the arm in which surveillance by urine tests was not more than 5% worse than results in the cystoscopy only arm, implying that the 2-year recurrence risk would not be less than 45% vs 50%. Statistical power was set to 80% and 1-sided significance was considered at 5%. Based on standard formulas for the required number of events in a log rank test 290 events had to be documented during followup. To evaluate the influence of knowledge of the urine test outcome the chi-square test would have 97% power to reveal differences in a recurrence rate of 30% vs 10% with 50 recurrences at a total of 250 followup visits with positive urine test results. Data were collected using standardized case record forms and analyzed by SPSS® 11.5 with $p < 5\%$ considered statistically significant.

RESULTS

Patients, Tumors and Urine Samples

Patients were recruited from July 2002 to June 2006 and followup ended by November 2006 (median followup 34 months). The 222 patients in the intervention arm (unblinded to urine test outcome) and the 226 controls (blinded to urine test outcome) were comparable in regard to gender, smoking pack-years, age distribution, histopathology and NMI UC status (primary vs recurrent) at study entry (table 1). A total of 3,379 cystoscopies were performed, including 1,501 in the intervention arm and 1,637 in the control arm. Of the 1,398 voided urine samples collected 1,073 (77%) were informative and in 325 (23%) the urine test failed due to various reasons.[10] Cystoscopy and corresponding urine test results were available for 837 followup visits, including 391