ACTAS
Urológicas Españolas

Predicting results of daily-practice cystoscopies

F. García-Velandria, J.F. Sánchez-García, L.A. Rodríguez-Toves,
L. Alvarez-Buitrago, C. Conde-Redondo, V. Rodríguez-Tesedo,
J.H. Amón-Sesmero, M. Cepeda-Delgado, A. Cobos-Carbó,
D. Alonso-Fernández, J.M. Martínez-Sagarra

Actas Urológicas Españolas

www.elsevier.es/actasuro
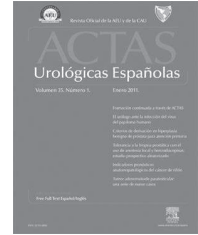
ORIGINAL ARTICLE

Predicting results of daily-practice cystoscopies ☆

F. García-Velandria[a], J.F. Sánchez-García[a], L.A. Rodríguez-Toves[a],
L. Alvarez-Buitrago[a], C. Conde-Redondo[a], V. Rodríguez-Tesedo[a],
J.H. Amón-Sesmero[a], M. Cepeda-Delgado[a], A. Cobos-Carbó[b,*],
D. Alonso-Fernández[a], J.M. Martínez-Sagarra[a]

[a] *Servicio de Urología, Hospital Universitario Río Hortega, Valladolid, Spain*
[b] *Unidad de Bioestadística, Departamento de Salud Pública, Universidad de Barcelona y Saalig Clinical, SL, Barcelona, Spain*

Abstract
Objective: Our objective was to elaborate a predictive model of bladder cancer, in an unselected clinical population submitted to cystoscopy.
Materials and methods: We recruited consecutive patients who underwent cystoscopy due to suspicion of bladder cancer or surveillance of a previously diagnosed bladder cancer. Urine cytology and a BTA-stat® (BTA) test were carried out for all patients. To avoid an assessment bias, the BTA-tests, cytologies and cystoscopies were conducted in a blinded fashion. We used logistic regression to predict cystoscopy results from cytology, BTA-test and clinical variables.
Results: From August 2011 to July 2012, we recruited 244 patients and 237 were valid for analysis. Newly diagnosed and surveillance cases were 13% and 87% respectively. Cytology and BTA-test sensitivities were 57.9% (CI 95: 42.2–72.1) and 63.2% (CI 95: 47.3–76.6) with specificities of 84.4% (CI 95: 78.7–88.8) and 82.9% (CI 95: 77.1–87.5). The predictive model included the BTA-test, cytology, time since previous tumor, and treatment with mitomicin or BGC during the last three months. The model predictive accuracy (AUC) was 0.85 (0.78–0.92), and dropped to 0.79 when excluding the BTA-test ($p = 0.026$). For the surveillance of bladder cancer, a 10% threshold on the model predicted probabilities resulted in an overall negative predictive value of 95.7%, and 95.0% in low grade tumors.
Conclusion: In a cost containment environment, our prediction model could be used to space out cystoscopies in patients with previous, low grade tumors, resulting in a more efficient use of resources in the healthcare system.
© 2013 AEU. Published by Elsevier España, S.L.U. All rights reserved.

PALABRAS CLAVE

Cistoscopia;
Técnicas Citológicas;
Neoplasias
genitourinarias;
Nomogramas;
Cáncer de vejiga
urinaria;
Marcadores tumorales

Predicción de resultados de cistoscopias en la práctica

Resumen

Objetivo: Elaborar un modelo predictivo de cáncer de vejiga en una población clínica no seleccionada derivada a cistoscopia.

Materiales y métodos: Pacientes consecutivos sometidos a cistoscopia debida a sospecha o seguimiento de un cáncer de vejiga previamente diagnosticado. Todos los pacientes fueron sometidos a citología urinaria y un BTA-stat®-test (BTA). Para evitar sesgos de evaluación, BTA, citologías y cistoscopias fueron realizados de forma ciega. Usamos regresión logística para predecir los resultados de la cistoscopia a partir de citología, BTA y variables clínicas.

Resultados: Entre agosto de 2011 y julio de 2012 seleccionamos 244 pacientes y 237 fueron válidos para el análisis. Un 13% fueron de nuevo diagnóstico y un 87% de seguimiento. Las sensibilidades de la citología y el BTA fueron 57,9% (IC 95%: 42,2–72,1) y 63,2% (IC 95%: 47,3–76,6) con especificidades de 84,4% (IC 95%: 78,7– 88,8) y 82,9% (IC 95%: 77,1–87,5). El modelo predictivo incluyó BTA, citología, tiempo transcurrido desde el diagnóstico del tumor previo y tratamiento con mitomicina o BGC en los últimos 3 meses. La precisión del modelo (AUC) fue 0,85 (0,78–0,92), y bajó a 0,79 al excluir el BTA (p = 0,026). En los casos de seguimiento, un umbral de 10% en las probabilidades predichas por el modelo resultó en un valor predictivo negativo de 95,7%, y 95,0% en los tumores de bajo grado.

Conclusión: En un contexto de contención de costes nuestro modelo puede usarse para espaciar las cistoscopias en pacientes con tumores de bajo grado previos, resultando en un uso más eficiente de recursos del sistema de salud.

Introduction

Bladder cancer is the most common malignancy of the urinary tract, with global age-standardized rates of 10.1 and 2.5 per 100,000 people/year (PY) for men and women respectively.[1] In Spain, the incidence of bladder cancer is 41.5 per 100,000 PY, with a mortality rate of 6.9 per 100,000 PY,[2] one of the highest figures in the European Union.

When bladder cancer is suspected, the initial assessment consists of urine cytology, cystoscopy, and a radiological study of the upper tract. Although at present urine cytology is the non-invasive test of choice, it has low sensitivity, especially in the case of low-grade tumors.[3] Many urine tests for the diagnosis of bladder cancer have been developed. Although different studies have shown the superiority of urine markers regarding sensitivity for bladder cancer detection as compared with cytology, several authors consider that none of these tests are ideal,[4] and that they are insufficient for effective monitoring and management of the patient.[5] In fact, current guidelines state that ''urinary cytology or markers cannot safely replace cystoscopy in the monitoring of non-muscle-invasive bladder cancer''.[1] However, other authors have suggested that urine tests may play a role in the monitoring of bladder cancer, replacing routine cytology in patients with low-grade/stage tumors[6] or extending the period between cystoscopies in the monitoring of patients with transitional cell carcinoma,[7,8] and facilitating the early detection of tumor recurrence in patients with high-grade tumors.[8]

The BTA stat test® (BTA) is a rapid, single-step immunochromatographic test for the detection of the antigen associated with bladder tumors in the urine,[4] which has significantly showed more sensitivity, and similar specificity when compared to cytology.[6] It is a test that is done in medical centers, with the characteristics of what an ideal test is considered: objectivity, accuracy, speed, and it is easily administered.[7] A study comparing 7 urine tests for bladder cancer in terms of validity (sensitivity and specificity), reliability (reproducibility), and efficiency (predictive value, complexity and costs) concluded that the most convenient (individual) assessment method was the BTA stat test®.[9]

Few attempts to develop a predictive model of bladder cancer have been made,[10,11] but none of them provided an appropriate model for routine clinical practice, since they were limited to selected patients or used a ''convenience'' sample. The aim of our study was to explore the possibility of developing such a predictive model of bladder cancer in an unselected and natural clinical population.

Patients and methods

We selected for our study all the consecutive patients who had undergone cystoscopy at our urology service throughout a year, either due to suspicion of bladder cancer or surveillance of previously diagnosed bladder cancer. The monitoring of previous tumors was carried out in accordance with the program recommended by the European guidelines.[1] Such guidelines were also followed to establish the indication for adjuvant treatment with MMC or BCG. Cytoscopies were always performed by the same team of experts in urology. Urine cytology and the BTA test were performed in all patients on the 7 days prior to cystoscopy. Urine samples for cytology (the second urine void in the morning) were taken on 3 consecutive days and cytology was always performed by the same team of general pathologists; the outcome was considered positive if malignant or atypical

cells were found in any of the 3 samples. The BTA test was performed on a spontaneous unrefrigerated urine sample, always on the 7 days after sample collection. In order to avoid a biased assessment, the evaluators of the cytology, cystoscopy, and the BTA test were blinded with regard to the results obtained in previous assessments.

The following variables were registered for all the selected patients: demographics (age, gender), previous diagnosis of bladder tumors (yes or no), time elapsed since the last diagnosed tumor, grade of the previous tumor (low or high, according to the 2004 WHO guidelines[12]), use of mitomycin or BCG in the last 3 months (yes or no) and the results of the cytology, BTA test and the cystoscopy (positive or negative).

Statistical methods

The measures of diagnostic accuracy (sensitivity, specificity and negative predictive values) and the corresponding 95% confidence intervals (CI 95%) were calculated on the basis of the cross-tabulation of the results of the cystoscopy (as standard reference), the BTA test and the cytology. The McNemar test was used to compare the BTA test and the cytology in terms of sensitivity and specificity. Logistic regression was used to investigate the capacity to predict the results of the cystoscopy on the basis of the results of the cytology and the BTA test, tumor grade, the time elapsed since the last diagnosed tumor (<18 months, 18 months or more, or no previous tumor) and the patient's age and gender. We used a retrograde elimination procedure to select the significant predictors ($p < 0.05$ in the Wald tests) and we internally validated the procedure to develop such models with the bootstrap method (10,000 replicas). For the selected model, we explored the possible interactions between markers (which were ignored if they were not significant), we traced the ROC curve for the probabilities derived from that model and we calculated the area under the curve (AUC) as a measure of predictive accuracy. We also explored two reduced models, excluding at the same time the cytology and the BTA test, to evaluate its relative contribution to the AUC. The effect of the prognostic factors of the model are described as *odd ratio* and the corresponding 95% confidence intervals (CI 95%) and a nomogram is provided to facilitate the use of this model in the clinical setting.

We explored the practical implications of the predictive model for monitoring patients with previous tumors using a threshold of 10% for the probabilities derived from the model, cross-tabulating them with the real results of the cystoscopy and estimating the measures of diagnostic accuracy for that threshold.

Data are described as n(%), mean (SD) or median (IQR) as appropriate. The statistical analysis was performed using R version 2.15.1.

Results

From August 2011 to July 2012, 244 patients were selected for our study, 237 of whom were valid patients for analysis (7 cases were excluded since values for crucial variables were missing). The patients' mean age (SD) was 65 (10) years and 191 of them were men. The breakdown of newly diagnosed

Table 1 Characteristics of the study patients.

Variable	Previous tumor (*n* = 206)	No previous tumor (*n* = 31)
Age[a]	65 (10)	63 (11)
Sex, male	167 (81.1)	24 (77.4)
Previous tumor		
Ca *in situ*	18 (8.7)	–
T1 GI	7 (3.4)	–
T1 GII	10 (4.8)	–
T1 GIII	28 (13.6)	–
Ta GI	40 (19.4)	–
Ta GII	74 (35.9)	–
Ta GIII	21 (10.2)	–
T2 GIII	1 (0.5)	–
Not available	7 (3.4)	–
Tumor grade, high	67 (33.5)	–
Mitomycin or BCG, last 3 months	55 (26.7)	–
Time since the last tumor (months)[b]	14 (29.7)	–
Result of the cytology, positive	39 (18.9)	14 (45.2)
Result of BTA, positive	46 (22.3)	12 (38.7)
Result of the cystoscopy, positive	23 (11.2)	15 (48.4)

The data are described as *n* (%).
[a] Mean (SD).
[b] Median (IQR).

cases and monitoring cases was 13% and 87% respectively, and their characteristics are described in Table 1.

Diagnostic accuracy of cytology and BTA test

The cytology and the BTA test detected 22 and 24 of the 38 positive cases of the cystoscopy, resulting in a sensitivity of 57.9% (CI 95%: 42.2–72.1) and 63.2% (CI 95: 47.3–76.6) respectively; and they were negative in 168 and 165 of the 199 negative cases of the cystoscopy, resulting in a specificity of 84.4% (CI 95%: 78.7–88.8) and 82.9% (CI 95: 77.1–87.5) respectively. The comparison between the cytology and the BTA test was not statistically significant either in terms of sensitivity (McNemar's chi-square = 0.1; df = 1; $p = 0.751$) or of specificity (McNemar's chi-square = 0.08; df = 1; $p = 0.770$). The negative predictive value was 91.3% (86.3–94.6) for the cytology and 92.2% (87.3–95.0) for the BTA test.

Prediction of the results of the cystoscopy

The result of the multivariate logistic regression analysis is shown in Table 2. Four variables were selected as independent predictive factors of a positive result of the cystoscopy. Tumor grade did not play a significant role once the above-mentioned 4 variables were taken into account. Bootstrap validation confirmed that the model shown in Table 2 was the most selected one among the 10,000 replicas of the

Table 2 Logistic regression model.

Sign	OR (95% CI)	p-Value
Independent term	0.24 (0.09–0.66)	0.005
BTA test	8.96 (3.37–23.83)	<0.001
Cytology	4.01 (1.66–9.67)	0.002
Mitomycin or BCG the last 3 months	0.21 (0.06–0.74)	0.014
Time since the last tumor[a]		
<18 months	0.31 (0.10–0.96)	0.042
18 months or more	0.13 (0.04–0.45)	0.001
ABC model = *0.854 (95% CI: 0.785–0.922)*		

CI: confidence interval; OR: *odds ratio*.
[a] Reference is no previous tumor.

backward elimination procedure. The interactions of the predictive factor were not statistically significant (results not shown).

When the model was reduced by excluding the result of the BTA test as a predictive factor, the AUC decreased from 0.854 (see Table 2) to 0.795 (CI: 0.710–0.880), which resulted in a significant reduction of the AUC ($p = 0.026$). When the BTA test was kept in the model, but the result of the cytology was excluded, the AUC was 0.839 (0.772–0.906) and this change was not statistically significant ($p = 0.344$). Fig. 1 shows the ROC curve for the 4-variable model and for the 2 reduced models which excluded the cytology or the BTA test. Fig. 2 presents a nomogram of the 4-variable model.

Implications of the model

Table 3 shows the result of patients with previous tumors classified by grade, the real result of the cystoscopy and the probability derived from the model using a threshold of 10%. When grouping low- and high-grade tumors, the overall sensitivity and specificity of the threshold of 10% was 73.9%
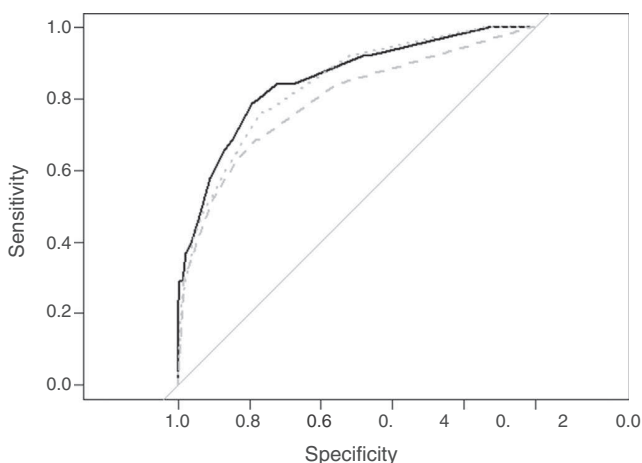

Figure 1 ROC curves of the predictive model (continuous line) of the four variables (see text) and two scale models, excluding the results of the BTA test (dashed line) or cytology (dotted line).
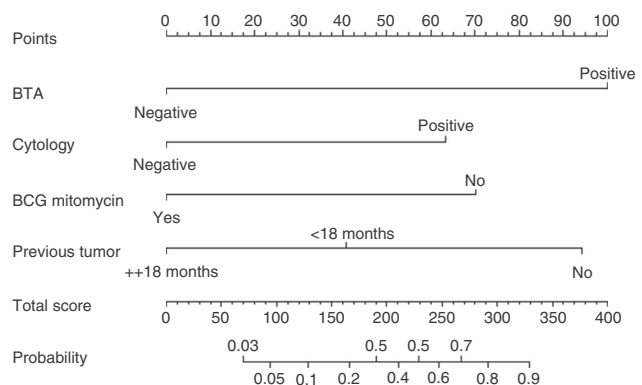

Figure 2 Nomogram for calculating the probability of a positive result in the cystoscopy from four predictors: BTA test, cytology, time since previous tumor, and treatment with mitomycin or BCG for the last 3 months.

(53.3–87.5) and 73.8% (67.0–79.6), with a negative predictive value (NPV) of 95.7% (91.0–98.0). This NPV implies that only 4.3% of the cases with a probability of 10% or less are positive in the cystoscopy. For low-grade tumors, the values of sensitivity, specificity, and NPV were 66.7% (41.7–84.8), 76.6% (68.4–83.2) and 95.0% (88.8–97.8) respectively. For high-grade tumors, those values were 87.5 (52.9–99.4), 67.8 (55.1–78.3) and 97.6 (87.4–99.9) respectively.

Discussion

Main results

Our results suggest that the BTA test is an independent predictive factor of bladder cancer, as well as cytology, the time elapsed since the last tumor and treatment with mitomycin or BCG in the last 3 months. Although tumor grade is a well-known predictive factor of recurrence in patients with previous tumors, it did not play a statistically significant role in our model, once the above-mentioned variables were taken into account. We believe that this is probably due to the strong association between the results of the BTA test and of the cytology and the result of the cystoscopy (OR 8.96 and 4.01 respectively, Table 2). For the patients with previous tumors the use of the estimated probabilities in our model with a threshold of 10% resulted in a remarkably high

Table 3 Probabilities of model and result of the cystoscopy.

Previous tumor grade (n = 206)	Result of the cystoscopy	Probability of the model[a]	
		<10%	≥10%
Low	Negative	95	29
	Positive	5	10
	Total	100	39
High	Negative	40	19
	Positive	1	7
	Total	41	26

[a] Probability of model of a positive cystoscopy.

negative predictive value (95.7% overall, 95% for low-grade tumors and 97.6% for high-grade tumors).

Despite showing similar sensitivity and specificity, the BTA test and the cytology appeared in our logistic regression analysis as independent predictive factors. This is due to the considerable lack of agreement when classifying patients as positive or negative ones. Thus, out of the 38 positive patients in the cystoscopy, 10 got discrepant results in the BTA test and in the cytology. Similarly, out of the 199 negative cases in the cystoscopy, 47 got discrepant results in the BTA test and in the cytology. Nevertheless, when we observed the discriminatory capacity (measured through the area under the ROC curve) of the models including, or excluding, the result of the cytology, it was not possible to find statistically significant differences ($p = 0.344$), but this may be due to low power.

Comparison with previous results

Previous studies have shown that the BTA test is more sensitive than the cytology,[4,6] and it shows similar specificity.[6] Our direct comparison of sensitivities lacked the capacity to detect that higher sensitivity of the BTA test, since the design of our study resulted in a small number of patients with bladder cancer ($n = 38$). With regard to specificity, our results are in line with previous research.

Shariat et al. developed a model to predict the recurrence and progression of the disease on the basis of a ''convenience'' sample of patients with transitional cell bladder carcinoma using a large, international cohort.[10] Their model included the patient's age and gender (although gender did not seem to contribute to model fitting, $p = 0.49$), the cytology and the nuclear matrix protein 22 test (NMP22) in the urine, with an overall predictive accuracy (AUC) of 0.798; this value is slightly lower than ours (0.85), maybe because they did not take into consideration the time elapsed since the last tumor and treatment with mitomycin or BCG, as we did. On the other hand, the patient's age was not selected in our backward elimination procedure, but since it remained at 25% of the 10,000 bootstrap replicas, we explored the impact of its inclusion in the model in terms of predictive accuracy and the increase of the AUC was not statistically significant (AUC 0.859, $p = 0.691$).

As for the newly diagnosed patients, Boman et al. presented a predictive accuracy (AUC) of 0.886 on the basis of a logistic regression model which included age, cytology and the BTA test.[11] However, their analysis was based on 3 selected groups (92 newly diagnosed bladder cancers, 64 cases of idiopathic microhematuria and 42 patients with nephritis) resulting in a prevalence of bladder cancer cases of 46.5%. This figure is disproportionately high for any unselected clinical population and may have enormously affected their results.

Huber et al. used logistic regression to evaluate the impact of several predictive factors on the result of the NMP22 test, instead of elaborating a predictive model of bladder cancer on the basis of the result of the test and other variables. Besides, they carried out their study in a cohort of retired employees who had been exposed to aromatic amines, and they only performed cystoscopies on those who had had a positive result in one of the 3 urine tests, probably

overestimating the sensitivity and the negative predictive value of the NMP22 test. For all these reasons their study is not comparable to ours.

According to our best knowledge, no other published study has tried to elaborate a predictive model of bladder cancer based on the current standard tests (cytology), on urine markers (such as the BTA test) and other clinical variables.

Constraints and advantages

A constraint of our study was the relatively low number of positive cystoscopies (38). This implies low power in the comparison of the sensitivity of the BTA test and the cytology. Likewise, for the estimations of logistic regression parameters to be reliable, a minimum number of 10 events per variable has been recommended.[13] Since our model has 4 variables, the number of events per variable is $38/4 = 9.5$, which is slightly below 10. Another obvious limitation of our study was the lack of external validation of our predictive model: the size of our sample prevented us from dividing it into two halves and, thus, using a half for the generation of the model and the other half for validation of the model. Besides, it has been proven that the discriminatory capacity of the models varies considerably among centers[10] and our study was carried out at a single center. Consequently, recreation in a subsequent multicenter study would benefit the confidence in our model. Finally, although the predictive accuracy of our model is high, it is far from perfect, and therefore cannot be used to replace cystoscopy.

Despite those limitations, we believe that our study has a number of merits: first of all, it provides evidence that the BTA test is an independent predictive factor of cytoscopy, not redundant of cytology, adding valuable information which results in a better predictive accuracy; secondly, we documented a threshold in the probabilities derived from the model, resulting in a considerably high negative predictive value (95.7%) for the monitoring of patients with previous tumors; and thirdly, the design of our study enabled a direct estimation of such predictive value, since we carried out our analysis in an unselected clinical population.

Interpretation

In a cost containment environment, our predictive model could be used to space out cystoscopies in patients with previous tumors. Using a threshold of 10% in the predicted probabilities of the model implies a negative predictive value of 95.7%; in our study this could have saved 141 cystoscopies (see Table 3) in a year, which is 59.5% of the total number of cystoscopies performed, resulting in a potential and significant improvement of the patients' comfort and in a more efficient use of resources in the health system. The nomogram shown in Fig. 2 was developed to facilitate the use of our predictive model in practice. It is worth noting that, with equal results in the BTA test and the cytology, those patients with no previous tumors will have a higher score than those patients with a previous tumor (resulting in a score above 90 points for no previous tumor, and a score of 70 for no treatment with MMC or BCG). This is due to the higher percentage of positive cystoscopies in patients with

no previous tumors than in patients with a previous tumor (48.4 *versus* 11.2%, from Table 1). Therefore, the nomogram will tend to produce higher probability estimates of positive cystoscopy in patients with no previous tumors.

Generalization

As noted above, there have been considerable variations in the predictive accuracy of the logistic regression models for bladder cancer. Several factors could explain these variations, such as differences in the casuistry, cytopathological diagnosis, treatment protocols, and patterns of recurrence, as well as the unmeasured variables and random variation. Therefore, our model should be validated, preferably in a multicenter study, to support a broad generalization.

Conflict of interest

Dr. García-Velandria reported no financial support from Sysmex Spain during the study and Dr. Cobos reports grants from Sysmex Spain, S.L., during the study. Drs. Sánchez-García, Rodríguez-Toves, Alvarez-Buitrago, Conde Redondo, Rodríguez-Tesedo, Amón-Sesmero, Cepeda-Delgado, Alonso-Fernández and Martínez-Sagarra have nothing to declare.

Acknowledgements

References

1. Babjuk M, Oosterlinck W, Sylvester R, Kaasinen E, Böhle A, Palou-Redorta J, et al. EAU guidelines on non-muscle-invasive urothelial carcinoma of the bladder, the 2011 update. Actas Urol Esp. 2012;36:389–402.
2. Campanario R, Saiz R, Carmona J, Romero E, Sánchez E, Castiñeiras J. Epidemiología cáncer de vejiga en España: estado actual, previsiones de futuro y comparación con los países de la Unión Europea. Vigo (España). In: LXXVII Congreso Nacional de Urología. 2012. p. 70.
3. Kaufman DS, Shipley WU, Feldman AS. Bladder cancer. Lancet. 2009;374:239–49.
4. Tilki D, Burger M, Dalbagni G, Grossman HB, Hakenberg OW, Palou J, et al. Urine markers for detection and surveillance of non-muscle-invasive bladder cancer. Eur Urol. 2011;60:484–92.
5. Frantzi M, Makridakis M, Vlahou A. Biomarkers for bladder cancer aggressiveness. Curr Opin Urol. 2012;22:390–6.
6. Lotan Y, Roehrborn CG. Sensitivity and specificity of commonly available bladder tumor markers versus cytology: results of a comprehensive literature review and meta-analyses. Urology. 2003;61:109–18.
7. Budman LI, Kassouf W, Steinberg JR. Biomarkers for detection and surveillance of bladder cancer. Can Urol Assoc J. 2008;2:212–21.
8. Lotan Y, Shariat SF, Schmitz-Dräger BJ, Sanchez-Carbayo M, Jankevicius F, Racioppi M, et al. Considerations on implementing diagnostic markers into clinical decision making in bladder cancer. Urol Oncol. 2010;28:441–8.
9. Sun Y, He D, Ma Q, Wan X, Zhu G, Li L, et al. Comparison of seven screening methods in the diagnosis of bladder cancer. Chin Med J (Engl). 2006;119:1763–71.
10. Shariat SF, Zippe C, Lüdecke G, Boman H, Sanchez-Carbayo M, Casella R, et al. Nomograms including nuclear matrix protein 22 for prediction of disease recurrence and progression in patients with Ta, T1 or CIS transitional cell carcinoma of the bladder. J Urol. 2005;173:1518–25.
11. Boman H, Hedelin H, Jacobsson S, Holmäng S. Newly diagnosed bladder cancer: the relationship of initial symptoms, degree of microhematuria and tumor marker status. J Urol. 2002;168:1955–9.
12. Sauter G, Algaba F, Amin M, Al E. In: Eble J, Sauter G, Epstein J, Sesterhenn I, editors. WHO classification of tumours of the urinary system and male genital organs. Lyon (France): IARCC Press; 2004.
13. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol. 1996;49:1373–9.