MODELLING OF SURVIVAL OF PATIENTS WITH COLON ADENOCARCINOMA BASED ON MULTIVARIABLE ANALYSIS OF THE STATE OF CANCER CELL NUCLEAR APPARATUS

A.N. Grabovoy[1,], O.O. Kolesnik[1], T.M. Savchyn[2], S.A. Antoniuk[1]*
[1]*National Cancer Institute, Kyiv 03022, Ukraine*
[2]*ESC "Institute of Biology" Taras Shevchenko National University, Kyiv 03022, Ukraine*

The aim: The creation of a mathematical model of survival in patients with colon adenocarcinoma based on multivariable analysis of the state of cancer cell nuclear apparatus. ***Patients and Methods***: The study was performed on 141 samples of biopsy materials or material obtained during surgical treatment of the patients with colon adenocarcinoma or benign colon neoplasms with the use of histological, morphometric, densitometric, immunohistochemical and mathematical methods. ***Results***: It has been shown that each discrete pattern of the state of adenocarcinoma cell nuclei (quantity of DNA, the number and volume of nuclear organizer regions, expression rates of Ki-67, Bcl-2 and p53) is prognostically invalid in the case of its separate use. Combination of these characteristics significantly enhances prognostic validity of the survival model. Based on equation of Cox proportional hazards, survival model of good quality for the patients with moderately and poorly differentiated adenocarcinoma and increased average DNA content in tumor cell nuclei has been created. ***Conclusion***: The proposed survival model for colon adenocarcinoma demonstrates the quality twice superior to the model based on the use of tumor grade only (G) which in fact is presently used as a sole common independent histological criterion of prognosis.
Key Words: colon adenocarcinoma, nuclear apparatus of tumor cells, survival.

Convenient systems for evaluation based on differentiation grade of colon adenocarcinoma (AC) sometimes are not optimal due to the difficulties of objective assessment and prognostic validity required for a choice of correct treatment of AC patients [1, 2]. Therefore, the search for new criteria and the development of prognostic means is of importance for clinicians [3–6].

In the process of neoplastic transformation nuclear apparatus undergoes significant changes caused by genome damage and its altered replication during cell division and realization of genetic information. In particular, structural organization and size of nuclei, state of nucleolar organizer regions, DNA quantity (polyploidy and aneuploidy), immunohistochemical markers of proliferation and apoptosis regulation, and a great variety of other features, have been studied. On large cohorts of AC patients there has been shown a relation of some patterns of the state of tumor cell nuclei with the properties of neoplasms and prognosis. However, when these schemes have been applied for prognosis in an individual patient, they were found to be of low significance with large error range. Therefore, discrete patterns could be considered as such which use for prognosis is of limited accuracy.

Some patterns of nuclear state are directly or indirectly interrelated, in particular, mitotic activity, DNA quantity (synthesis, ploidy, aneuploidy), the state of nuclear organizer regions (NOR), apoptosis regulation. Increased and uncontrolled mitotic activity is among the main features of tumor growth. This could be accompanied with malfunction of mitosis mechanisms (upon chromosomal instability) leading to the development of polyploidy or aneuploidy. Along with this, DNA synthesis preceding division, at conditions of high proliferation activity results also in a certain increase of total DNA quantity in tumor cells. From other side, lesion of mechanisms of apoptosis and survival of the cells, especially the cells with increased content of DNA in nuclei, also promotes an elevation of medium DNA content in tumor cell nuclei.

The aim of the work: the creation of a mathematical model of survival in patients with colon AC based on multivariable analysis of the state of cancer cell nuclear apparatus.

PATIENTS AND METHODS

The study was performed on 141 samples of biopsy materials or materials obtained during surgical treatment of the patients with colon AC or benign colon neoplasms: conventionally normal colon tissue (C; n = 11); polyps and adenomas (B; n = 16); AC of different grades: G1 (n = 28); G2 (n = 73); G3 (n = 13).

The material was fixed in a buffered 10% formaline (pH 7.4, 4 °C, 24 h) and embedded in paraplast with the use of Histos-5 histoprocessor (Milestone, Italy). From paraffin blocks, 5 μm sections were cut using Microm HM325 (Thermo Scientific, Germany). The sections were stained with hematoxylin and eosin for common tumor assessment, or with Einarson's gallocyanin chrome alum stain (pH 1.62, 37 °C, 24 h) for determination of nucleic acid content [7]. In each case, a part of sections was treated with RNAase (Macherey-Nagel GmbH & Co. KG, Germany) for RNA destruction [8]. NOR were detected using silver nitrate impregnation. Immunohistochemical reactions were performed with

the use of mouse monoclonal antibodies against antigen Ki-67 (MIB-1, Dako, Denmark), Bcl-2 oncoprotein (Dako, Denmark), and p53 protein (Dako, Denmark) according to the instructions of the manufacturer and visualized using detection system EnVision™ FLEX (Dako, Denmark). The sections were additionally stained with Gill hematoxylin.

The preparations were examined and photographed using microscope Nikon Eclipse 80i supplied with camera DS-5SMc/L2 at standardized conditions, images (× 400, 1280 × 960 pixels RGB) were processed using analysis system ImageJ 1.46. In images of the preparations stained with gallocyanin chrome alums, in each 30 tumor cells there have been determined cross-sectional area of cell nucleus, medium (DM) and integral optical density of nucleus, and the quantity of nuclear DNA (NDNA). For calculation of the latest index, DNA content in lymphocyte nuclei was accepted as a unit (c.u.) [9]. The tumors were distributed into three ranks by average DNA content: D — average DNA content up to 1.2 c.u., D+ — average DNA content from 1.2 to 2.5 c.u., T+ — average DNA content higher than 2.5 c.u. [10]. In silver-impregnated sections, in each 50 tumor cells there was determined cross-sectional area of nucleus, quantity of NOR in nucleus (nNOR), and diameter of each NOR with following calculation of their total volume (vNOR). In each tumor, the percent of cells expressing Ki-67, Bcl-2, and p53 has been determined.

Statistical analysis was performed with the use of IBM SPSS Statistics 22. For assessment of normality, one sample Kolmogorov — Smirnov test was used. Correlation analysis was performed using Spearman rank correlation method. The differences between the groups were evaluated by Mann — Whitney U test. Survival analysis was performed by Kaplan — Meier estimator using log-rank test for evaluation of the differences between the groups.

For creation of survival models, Cox proportional hazards model (Cox regression model) was used. Standard form of regression equation is:

$$S(t) = S_0(t)^{e^{(\beta_1 x_1 + \ldots + \beta_i x_i)}},$$

where: $S(t)$ — survival at time t; $S_0(t)$ — baseline survival; β_i — regression parameters; x_i — the studied characteristics of tumor cell nuclear apparatus: X_1 — 0 if AC differentiation grade is G2, 1 — if AC differentiation grade is G3; X_2 — ratio of nNOR/Ki-67 values; X_3 — product of nNOR × NDNA values; X_4 — derivative of nNOR × NDNA/Ki-67; X_5 — DNA content rank, may be of three values (D, D+ or T+); X_6 — 0 corresponds to rank D+, 1 — to rank T+.

To build the models of prognosis on the basis of Cox regression, different number of nuclear apparatus features were used.

Model I includes just a single characteristic X_1.

Model II includes X_2, X_3, and X_4 (three characteristics).

Model III includes all 4 characteristics of models I and II.

Model IV includes 5 characteristics: X_1, X_2, X_3, X_4, and X_5.

Model V also includes 5 characteristics: X_1, X_2, X_3, X_4, X_6.

The prognostic validity of the developed models was compared with that of basic Cox model that doesn't include any characteristics of tumor process with the use of likelihood ratio test (LR). As an index of prognosis model quality (survival of patients) there has been used ROC-analysis where area under the curve (AUC) reflects model quality as follows: 0.5–0.6 — bad, 0.6–0.7 — fair, 0.7–0.8 — good, 0.8–0.9 — very good, 0.9–1.0 — perfect. The differences were considered significant if p < 0.05.

RESULTS AND DISCUSSION

The performed study has shown that average DNA content in cell nuclei of tumors of different grade varied from 2.06 (G3) to 2.38 (G1) (Table 1). While the differences between groups C and B were significant (p < 0.05), the differences between various G were insignificant (p > 0.05). In the studied sample among AC of different grades (B, G1, G2, G3) there were the tumors with different DNA content ranks (D, D+, T+).

Table 1. Quantitative values of features of nuclear apparatus of tumor cells of colon AC and benign colon neoplasms

Feature	Conventionally normal colon tissue		Benign colon neoplasms		Colon AC					
					G1		G2		G3	
	M	m	M	m	M	m	M	m	M	m
NDNA, cu	1.32	0.07	2.22*	0.31	2.38	0.20	2.06	0.09	2.06	0.16
nNOR, n	1.70	0.07	3.02**	0.13	2.90	0.16	2.39**	0.09	2.25	0.19
D	1.87	0.03	2.77	0.05	2.85	0.19	2.19	0.31		
D+	1.66	0.08	2.74**	0.15	3.09	0.32	2.46*	0.11	2.36	0.22
T+			3.47	0.21	2.73*	0.23	2.30	0.22	2.01	0.38
vNOR, mkm³	1.96	0.08	2.80**	0.18	2.65	0.19	2.86	0.14	3.25	0.43
D	1.93	0.01	2.58	0.20	2.51	0.13	2.45	0.45		
D+	1.97	0.10	2.56*	0.21	2.70	0.35	2.89	0.15	2.86	0.42
T+			3.18	0.36	2.66	0.29	2.98	0.36	4.14	1.00
Ki-67, %	6.16	3.79	34.56*	6.37	36.59	5.00	46.07	3.22	49.09	8.50
D			15.96	7.45	41.14	6.61	54.80	9.94		
D+			37.66*	8.47	27.43	7.01	46.81*	4.05	50.58	9.39
T+			37.16	12.36	44.52	9.42	39.44	6.16	45.75	20.12
Bcl-2, %	28.95	3.87	33.88**	10.26	41.23	8.27	40.73	5.28	17.60	7.48
D			44.00	22.48	30.80	19.45	44.42	15.98		
D+			39.71	18.93	32.63	12.31	40.96	6.56	24.15	10.16
T+			22.00	13.92	55.36	13.65	37.82	11.32	2.85	1.77
p53, %	6.28	3.54	14.95**	7.60	50.50**	7.52	65.70*	4.59	37.43*	10.21
D			12.40	12.40	50.00	20.86	54.46	14.03		
D+			0	0	41.98*	11.12	63.34*	5.90	38.83	12.29
T+			33.67	17.55	60.01	12.01	80.16	7.03	34.27*	21.03

Note: NDNA — content of DNA in nucleus (cu); D, D+, T+ — DNA content ranks; nNOR — number of nucleolar organizers; vNOR — total volume of nucleolar organizers; Ki-67, Bcl-2, p53 — expression level of corresponding marker in tumor; p — significance: *p < 0.05, **p < 0.01 (comparing with previous group).

If one takes into account the variability of this index, its use as a criterion of tumor grading and independent prognostic factor seems to be improper [9, 11]. Also, the decrease of DNA content ranks in nuclei of tumor cells along with G increase in combination with restriction of cellular heterogeneity via reduction on the number of hyperploid cells [9], supports the realization of clonal evolution during AC progression [8, 12].

Number of NOR in tumor cell nuclei tended to decrease along with increase of grade while their total volume (vNOR) has been increasing; however, the differences between these indexes in approximal tumor groups were insignificant. Therein, the direct dependence between nNOR/vNOR was not found. An absence of proportional increase of nNOR/vNOR in tumor cell nuclei along with elevation of DNA content is evidence of DNA dysfunction or inactivity [9]. This phenomenon could be considered as a manifestation of decreased viability of these cells leading to ageing, apoptosis, and elimination of such morphofunctional types from a tumor [13, 14].

All studied tumors have shown Ki-67 expression index higher than zero, which ranged from 2.17% to practically 100% with a median of 37%. In general, there was recorded a sharp rise of tumor cell Ki-67 expression index in group B *vs.* C with its following moderate increase. Observed tumors have wide range of Bcl-2 and p53 expression. The correlation between G and DNA content in nucleus of tumor cell was noticed.

The performed correlation analysis has shown that there are a number of statistically significant weak relations between the parameters. G correlated with nNOR ($r = -0.340$; $p < 0.01$) and their average volume ($r = -0.348$; $p < 0.01$), but had no correlation with vNOR. Apart from this, there was recorded a relation between increase of Ki-67 labeling index and tumor grade ($r = 0.193$; $p < 0.05$), and also with nNOR ($r = -0.202$; $p < 0.05$). Positive correlation was determined also for p53 and G ($r = 0.227$; $p < 0.05$).

Survival analysis dependent on discrete patterns has shown that it significantly differed only in the groups distributed only by grade (G) ($p < 0.01$). No statistically significant differences between Bcl-2(+/−) and p53(+/−) groups has been revealed (Table 2). However, in the tumors with different DNA content ranks there has been observed a tendency for survival change. It was the highest in rank D and the lowest in rank T+. This fact could be explained by the different mechanisms of development of these neoplasms [15]. Tumors from rank D may be refer to ones which develop via epigenetic mechanisms and microsatellite, while ranks D+ and T+ — chromosome instability.

The data obtained in correlation analysis and survival analysis in the groups allowed to build a survival model using Cox regression (proportional hazards model). Initially the models have been created and there was performed an analysis by the method of direct inclusion of large factor array for their assessment as prognostic criteria for all sampling, including benign neoplasms as well (Table 3). The data of regression analysis (Cox regression model) have demonstrated that the use of factors which characterize the state of tumor cell nuclei discretely or independently of each other, did not allow to build a significant model. This is related to insignificant ($p > 0.05$) influence of these factors on survival, and their impact did not alter significantly regression equation.

Table 2. Comparing of patients survival with colon AC in different groups using log-rank test (null hypothesis – there is no difference between groups)

Distribution of tumors into subgroups by the patterns	χ^2	p
G (B, G1, G2, G3)	42.617	<0.01
Quantity of DNA (D, D+, T+)	1.638	0.441
Bcl-2+ and Bcl-2−	0.054	0.816
p53+ and p53−	0.116	0.733
Ki-67 (distribution relative to median (37%))	0.275	0.600
nNOR (distribution relative to median (2.57))	0.181	0.671

Note: G – tumor grade; D, D+, T+ – DNA content ranks; Bcl-2 – Bcl-2 expression in tumor; p53 – expression in tumor.

Table 3. Prognostic significance of discrete features of nuclear apparatus of tumor cells (Cox model)

Feature	β	SE	p
Tumor grade (G)	2.097	0.432	< 0.01
vNOR	0.128	0.202	0.526
Ki-67, %	0.001	0.009	0.897
Bcl-2, %	0.002	0.006	0.731
p53, %	0	0.006	0.98
NDNA	0.103	0.248	0.678
nNOR	0.065	0.278	0.814
D, D+, T+	0.452	0.375	0.228

Note: β – coefficient; SE – standard error of coefficient; p – significance; nNOR – number of nucleolar organizers; vNOR – total volume of nucleolar organizers; Ki-67, Bcl-2, p53 – expression level of corresponding marker in tumor; NDNA – quantity of DNA cu in nucleus; DM – medium optical density of nucleus; G – tumor grade; D, D+, T+ – DNA content ranks.

As far as tumor grading (G) is the only one significant prognostic criterion, it has been used in the following analysis as a base for model building. Apart from this, to elevate model accuracy groups B and G1 were excluded because in these groups there has been recorded 100% 5-year survival rate, what preconditions their uninformativeness.

For assessment of prognostic validity of a combination of the factors, we have performed a number of procedures of direct and inverse inclusion of the variables in regression model. There have been used either discrete factors or their combinations and their interrelation has been analyzed (Table 4).

Inclusion of two tumor grading values — G2 and G3 into regression equation produced a significant model (Model I), but with low AUC (0.339) (Figure). However, positive value β = 1.87 indicates that the higher G increases the likelihood of lethal outcome.

The search for significant combination of factors allowed to include such elements as NDNA, nNOR and Ki-67 which formed a base for Model II. Its analysis has shown that simultaneous increase of NDNA and nNOR increases the risk of death. However, increase of nNOR in parallel with the decrease of Ki-67 expression index elevates the risk. The quality of this model is practically equal from the first one what indicates their similarly low prognostic validity (see Table 4).

Joining of Models I and II into a new Model III has been found to be rational because it resulted in a decrease of LR and an increase of AUC up to 0.605 (see Figure),

what however couldn't be considered statistically significant at p = 0.18. Better result was obtained after inclusion of feature of DNA content ranks (D, D+, T+) into regression equation. This procedure led to the model of better quality (Model IV): AUC = 0.663 (see Table 4). Further, the best model was obtained using D+ and T+ ranks (Model V), AUC = 0.722 — fair quality. Thus, Model V based on 5 features: X_1, X_2, X_3, X_4, X_6, acquired the best properties of quality and informativeness among other models.

Table 4. Regression parameters and features included to the model upon regression analysis of survival (Cox model) of AC patients

Model	n	LR	p (LR)	AUC	p (AUC)	Parameter/s	β	SE	p (β)
I	74	10.42	<0.01	0.339	0.05	X_1	1.870	0.504	<0.01
II	74	8.95	0.03	0.391	0.18	X_2	0.159	0.063	0.01
						X_3	0.239	0.080	<0.01
						X_4	−0.069	0.028	0.01
III	74	20.67	<0.01	0.605	0.18	X_1	2.031	0.517	<0.01
						X_2	0.173	0.066	<0.01
						X_3	0.265	0.086	<0.01
						X_4	−0.074	0.028	0.01
IV	74	25.44	<0.01	0.663	0.04	X_1	1.932	0.531	<0.01
						X_2	0.279	0.087	<0.01
						X_3	0.211	0.094	0.03
						X_4	−0.118	0.038	<0.01
						X_5	1.459	0.692	0.04
V	65	28.14	<0.01	0.722	<0.01	X_1	1.936	0.557	<0.01
						X_2	0.378	0.107	<0.01
						X_3	0.232	0.104	0.03
						X_4	−0.159	0.049	<0.01
						X_6	2.006	0.801	0.01

Note: n — number of cases included into analysis; LR — likelihood ratio (comparing with null model); AUC — area under ROC curve; β — coefficient; SE — standard error of coefficient; p ( ) — significance for LR, AUC and β; X_1 — tumor grade (G2 or G3); X_2 = nNOR/Ki-67; X_3 = nNOR × NDNA; X_4 = nNOR × NDNA/Ki-67; X_5 — DNA content ranks (D, D+ or T+); X_6 — DNA content ranks (D+ or T+).
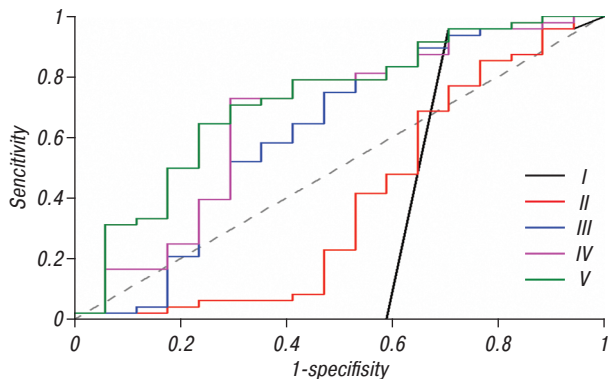

Figure. ROC-analysis of different survival models in colon AC. I–V — ROC curves for different models

$S_0(t)$ value is calculated on the basis of sampling by Cox regression with the use of SPSS program for different t values (Table 5). Using the baseline function values at time t, and the values of equation variables (nNOR, Ki-67, NDNA, G) obtained by their direct measurement in an individual patient, and introducing them into formula 1 accounting the values of the coefficients, the values of survival function for the patient $S(t)$ for different t could be calculated (see Table 5).

So, the performed studies have shown that discrete patterns of AC cell nuclei state (DNA quantity, number and volume of NOR, expression rates of Ki-67, Bcl-2 and

p53) are prognostically invalid in a case of their separate use and could not be considered as direct prognostic criteria. However, the derivatives of a number of patterns could serve as significant criteria for prognosis (in particular, nNOR/Ki-67, nNOR × NDNA, and nNOR × NDNA/Ki-67). Besides, upon their use in survival model mathematically expressed dependence is not linear and is described by more complex functions (in this case, exponential one). Good quality survival model based on Cox proportional hazards model for patients with AC of G2 and G3, with increased DNA content in tumor cell nuclei, has been created.

Table 5. Estimated survival function for patients with colon AC (Model V) for mean values of features

t	3	5	7	8	9	10	11	36
$S_0(t)$	1.000	0.998	0.996	0.994	0.993	0.992	0.985	0.980
$S(t)$	0.996	0.982	0.968	0.951	0.941	0.931	0.876	0.836

Note: t — survival time, months; $S_0(t)$ — baseline survival at time t; $S(t)$ — survival function at time t.

The proposed survival model in the case of AC demonstrates the quality twice superior to that of a model based just on the use of tumor grade (G), which presently is a single common independent histological criterion of prognosis [3, 2].

REFERENCES

1. Compton CC, Fielding PL, Burgart LJ, *et al*. Prognostic factors in colorectal cancer. Arch Pathol Lab Med 2000; **124**: 979–94.

2. WHO classification of tumors the digestive system. Lyon: IARC, 2010: 131–81.

3. Barresi V, Bonetti LR, Ieni A, *et al*. Histological grading in colorectal cancer: new insights and perspectives. Histol Histopathol 2015; **30**: 1059–67.

4. Bolocan A, Ion D, Ciocan D, *et al*. Prognostic and predictive factors in colorectal cancer. Chirurgia (Bucur) 2012; **107**: 555–63.

5. Coppedè F, Lopomo A, Spisni R, *et al*. Genetic and epigenetic biomarkers for diagnosis, prognosis and treatment of colorectal cancer. World J Gastroenterol 2014; **20**: 943–56.

6. Kulendran M, Stebbing J, Marks C, *et al*. Predictive and prognostic factors in colorectal cancer: a personalized approach. Cancers (Basel) 2011; **29**: 1622–38.

7. Luppa Kh. Basic histochemistry. Moscow: Myr, 1980. 344 p. (in Russian).

8. Sottoriva A, Kang H, Ma Z, *et al*. A Big Bang model of human colorectal tumor growth. Nat Genet 2015; **47**: 209–16.

9. Grabovoy AN, Antoniuk SA, Vorobey EA. Nucleic acids contents in the colon epithelial tumor cells nuclei. Ukr Morfolohichnyi Almanakh 2013; **11**: 73–6 (in Ukrainian).

10. Avtandylov HH. Diagnostic medical ploidometry. Moscow: Medytsyna, 2006. 192 p. (in Russian).

11. Kørner H. Microsatellite instability and DNA ploidy in colorectal cancer. Cancer 2009; **115**: 271–82.

12. Ding L, Raphael BJ, Chen F, Wendl MC. Advances for studying clonal evolution in cancer. Cancer Lett 2013; **340**: 212–9.

13. Davoli T, de Lange T. The causes and consequences of polyploidy in normal development and cancer. Ann Rev Cell Dev Biol 2011; **27**: 585–610.

14. Holland AJ, Cleveland DW. Losing balance: the origin and impact of aneuploidy in cancer. EMBO 2012; **13**: 501–14.

15. Haigis K. Molecular pathogenesis of colorectal cancer. Spr Science Business Med 2013; **6**: 320.