

A NEW METHOD TO DIAGNOSE CANCER BASED ON IMAGE ANALYSIS OF MASS CHROMATOGRAMS OF VOLATILE ORGANIC COMPOUNDS IN URINE

UDC 61–006–074

Received 11.09.2014



A.A. Rozhentsov, D.Tech.Sc., Professor, Head of the Department of Radiotechnical and Biomedical Systems¹;

A.V. Koptina, PhD, Doctoral Candidate¹;

A.A. Mitrakov, Postgraduate, the Department of Radiotechnical and Biomedical Systems¹;

Head of the Department of Endoscopy²;

T. Sharipova, Student¹;

I. Tsapaev, Student¹;

V.L. Ryzhkov, Postgraduate, the Department of Radiotechnical and Biomedical Systems¹; Endoscopist³;

K.A. Lychagin, Postgraduate, the Department of Radiotechnical and Biomedical Systems¹;

R.R. Furina, Postgraduate, the Department of Radiotechnical and Biomedical Systems¹; Endoscopist³;

N.N. Mitrakova, D.Med.Sc., Professor, the Department of Radiotechnical and Biomedical Systems¹;

Head of the Department of Endoscopy³

¹Volga State University of Technology, Lenin Square 3, Yoshkar Ola, the Mari-El Republic, Russian Federation, 24000;

²Nizhny Novgorod Regional Oncologic Dispensary, Branch No.1, Rodionova St., 190, bld. 5, Nizhny Novgorod, Russian Federation, 603126;

³Republic Clinical Hospital, Osipenko St., 33, Yoshkar Ola, the Mari-El Republic, Russian Federation, 424000

The aim of investigation was to develop a technique to diagnose early cancer based on the image analysis of mass chromatograms of volatile organic compounds in urine, and assess its efficacy.

Materials and Methods. The patients were grouped by nosological types: lung cancer, esophageal cancer, gastric carcinoma, colon cancer.

A control group consisted of healthy individuals. Metabolic profiles of volatile organic compounds of urine were obtained using head space solid phase microextraction coupled to gas chromatography — mass spectrometry method.

Results. In accordance with the processed data of mass chromatograms of volatile organic compounds of urine for each patient and their comparison with images of control individuals, we made reference metabolic profiles of oncological diseases. The proposed technique for cancer detection is easy, non-invasive, low labor intensive and affordable. According to the assessment of the results obtained, the method sensitivity is 100%, specificity — 90.62%, type I error probability — 0%, type II error probability — 9.38%.

Conclusion. The findings could be used for the further understanding of etiology and pathology of various forms of oncological diseases.

Key words: urine volatile organic compounds; cancer biomarker; gas chromatography — mass spectrometry; metabolic profile.

Early diagnostics is the necessary condition of effective treatment of oncological diseases [1]. Currently, there is a great number of approaches to solving the problem; however, screening studies based on these approaches are not always available due to no access to qualified medical care, financial and organizational problems, low capacity of the equipment used [2–5]. In this regard, currently, it is a critical task to develop a diagnostic technique for cancer diseases. The technique is to have low labor intensity and require no costs for carrying out the analysis providing means for screening programs.

To make metabolic profiles of oncological diseases can be one of the approaches [6–8]. The most available information source of human metabolism is urine, in particular, its volatile fractions [9–12]. Urine sampling and analysis is a routine procedure in preventive medical examination of population, therefore, to carry out an extra test presents no organizational problems. Thus,

the construction of metabolic profiles of oncological diseases based on urine volatile organic compounds is of scientific interest.

The aim of investigation was to develop an early technique to diagnose cancer based on the image analysis of mass chromatogram of volatile organic compounds in urine, and assess its efficacy.

Materials and Methods. The patients with confirmed cancer types were grouped to reveal metabolic profiles of oncological diseases: gastric carcinoma (n=12), lung cancer (n=14), colon cancer (n=8), esophageal cancer (n=12). The selection of these cancer types is due to the fact that in Russia they rank first in the mortality rate: lung cancer — 18.5%, gastric carcinoma — 13.5%, and colon cancer — 12.7% [1]. In addition, there was a control group consisting of healthy volunteers (n=35).

The present study complies with the declaration of Helsinki adopted in June, 1964 (Helsinki, Finland) and

For contacts: Mitrakova Nina Nicolaevna, e-mail: endomitrakova@mail.ru

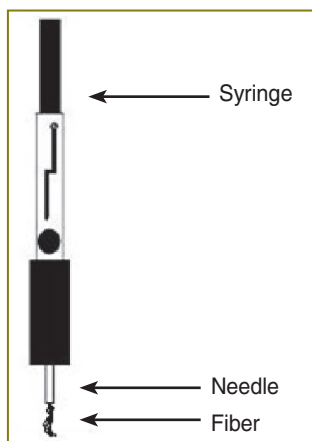


Fig. 1. A TRME syringe

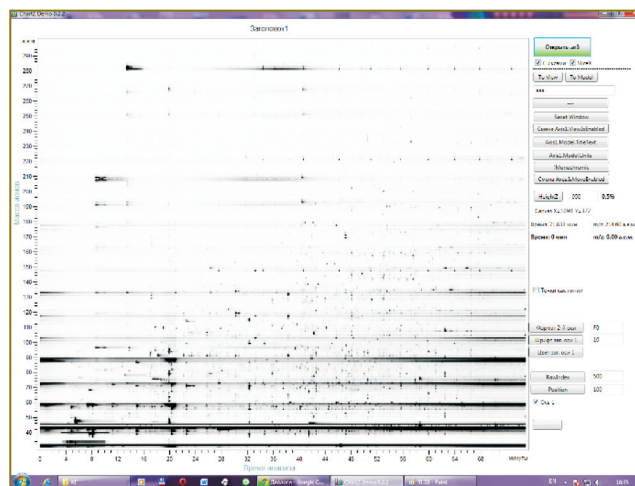
revised in October, 2000 (Edinburg, Scotland), and was approved by the Ethics Committee of Volga State University of Technology. All patients gave written informed consent.

The compositional analysis of urine volatile organic compounds was carried out using chromatato-mass-spectrometry. For sample preparation we used three-phase micro-extraction (TPME), the samples of volatile organic compounds being withdrawn by means of special fiber in a syringe (Fig. 1).

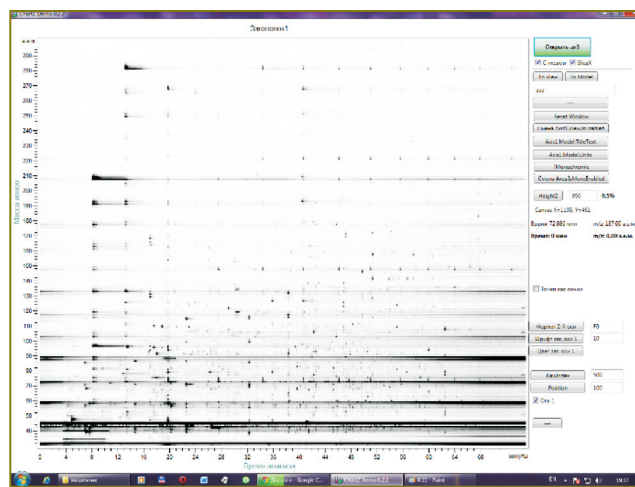
Before each experiment a needle was condensed to remove the previous sample residuals; to accomplish this, the needle was put in a vaporizing device at 250°C, assigning the following mode of chromatograph operation: column temperature — 250°C; carrier gas (helium) storing through a column — 1 ml/min; vaporizing device temperature — 250°C.

After 5-minute condensation the fiber was drawn in a needle, the syringe being withdrawn. Then, 15 ml of patient’s urine was put in a special hermetic and sterile tube. For better evaporation of urine volatile fractions, the sample was successively added by 250 µl of concentrated HCl and 0.8 g of NaCl salt. The obtained solution was shaken in such a way that the salt was totally dissolved, with foam free. After that the tube was placed in a constant-temperature bath at 48°C. TRME fiber was put in a tube above the urine sample followed by the absorption of volatile substances for 15 min. Then the fiber was drawn in a needle, withdrawn from the tube and placed in the chromatograph evaporator, where volatile substances were desorbed and separated into components. In a mass-spectrometer detector the components were separated into ions and recorded forming a mass-spectrum. Six minutes after examination mode engagement the syringe was taken out of the vaporizing device. The analysis was completed 75 min later resulting in a mass chromatogram representing the time-based sweep of a mass-spectrometer detector output signal. Fig.2 shows an example of mass chromatogram images. A detector output signal level was coded by image brightness.

A standard processing procedure of mass chromatograms is the identification of substances corresponding to the peaks marked in a chromatogram by comparing the obtained and the reference mass-spectra. We excluded from the list the substances released from the column surface and other chromatograph components (Table 1).



a



b

Fig. 2. Images of 3D-spectrograms: a — control group; b — colorectal cancer

Table 1
The list of substances excluded

No.	Substances
1	Cyclotetrasiloxane, octamethyl-
2	Cyclotrisiloxane, hexamethyl-
3	Disiloxane, hexamethyl-
4	Heptasiloxane, hexadecamethyl-
5	Hexasiloxane, tetradecamethyl-
6	Trisiloxane, 1,1,1,5,5,5-hexamethyl-3,3-bis(trimethylsilyl)oxy-

Table 2
Metabolic profiles of different groups

Control group	Gastric carcinoma	Lung cancer	Colon cancer
Benzene, 4-ethenyl-1,2-dimethyl-	2,6-Bis(1,1-di-methylethyl)-4-(1oxopropyl)phenol	2H-1-Benzopyran, 3,4,4a,5,6,8a-hexahydro-2,5,5,8a-tetramethyl-(2 α ,4 α ,8 α)-	Furan, 2,4-dimethyl-
Bicyclo[3,1,0]hex-2-ene, 4-methylene-1-(1-methylethyl)-	7-Octen-2-ol, 2,6-dimethyl-	Naphthalene, 1,2,3,4-tetrahydro-1,1,6-trimethyl-	Naphthalene, 1,6,7-trimethyl-
1,2-Bis(trimethylsilyl)-4,5-dimethylcyclohexa-1,4-diene	5-Hepten-2-one, 6-methyl-	2-Buten-1-one, 1-(2,6,6-trimethyl-1,3-cyclohexadien-1-yl)-, (E)-	3-Cyclohexene-1-methanol, α , α ,4-trimethyl-, (S)-
Acetamide	2-Cyclohexen-1-one, 3-(2,3,4,9-tetrahydro-1H-pyrido[3,4-b]indol-1-yl)methyl-	1,5,5-Trimethyl-6-(3-methyl-butyl-1,3-dienyl)-cyclohexene	β -Pinene
Phenyl-pentamethyl-disiloxane	2-Fluorobenzoic acid, 4-nitrophenyl ester	Methylene Chloride	3-Buten-2-ol, 2-methyl-
Ethane, 1,2-dichloro-	6-Methyl-6-(5-methylfuran-2-yl)heptan-2-one	2-Butenal, 3-methyl-	1,3-Cyclopentadiene, 5-(1-methylethylidene)-
Phosphorothioic acid, O-(2,6-dichloro-4-methylphenyl) O,O-dimethyl ester	Furan, 2-methyl-5-(methylthio)-	2-Hexanone, 5-methyl-	1,3,6-Octatriene, 3,7-dimethyl-, (E)-
Naphthalene, 1,2-dihydro-4-methyl-	Ethanol	Phenol, 3,5-dimethyl-	Hexanoic acid, 2-methyl-
Methyl Isobutyl Ketone	Ethanol, 2-phenoxy-	(R)-(+)-3-Methylcyclopentanone	Benzene, 3-(2-cyclohexylethyl)-6-cyclopentylhexyl-
Butanoic acid, 2-methyl-	Glycine, N-4-(trimethylsilyloxybenzoyl-, methyl ester	α -Methylstyrene	Pentanoic acid, 2,2,4-trimethyl-3-carboxyisopropyl, isobutyl ester
o-Xylene	Naphthalene, 1-methyl-7-(1-methylethyl)-	Glycine, N-phenyl-, ethyl ester	1-Methyl-1-(2-hydroxyethyl)-1-silacyclobutane
Phenol, 4-chloro-2,6-bis(1,1-dimethylethyl)-	Acetic acid, 2-ethylhexyl ester	5,7a-Didehydroindicine pertrimethylsilyl ether	Vanillin, tert-butyl dimethylsilyl ether
1-Boraindane, 3-methyl-1-1-(trimethylsilyl)-2,4-cyclopentadien-3-yl-	1H-Pyrazole, 4,5-dihydro-5,5-dimethyl-4-isopropylidene-	2-Cyclopenten-1-one, 3-methyl-	Acetamide, N,N'-ethylenebis(N-nitro-
3-Ethylcyclopentanone	1H-Indene, 2,3-dihydro-1,1,3-trimethyl-3-phenyl-	Ethanone, 1-(2-hydroxyphenyl)-	Benzene, (dibromomethyl)-
Benzene, 1,3,5-trimethyl-	3-tert-Butylsulfanyl-3-fluoro-2-trifluoromethyl-acrylic acid methyl ester		
1,1,6,8-Tetramethyl-1,2-dihydro-naphthalene	Benzoic acid, 2-ethylhexyl ester		
Benzonitrile	1H-Inden-1-one, 2,3-dihydro-3,3,5,6-tetramethyl-		
Ethylbenzene	Cyclohexane, 1,3,5-trimethyl-2-octadecyl-		
2-Nonanone	Naphthalene, 2,3,6-trimethyl-		
1-Bromo-2-methyl-2propanol			
4-Methyl-2,4-bis(4'-trimethylsilyloxyphenyl)pentene-1			
Ethanedioic acid, bis(trimethylsilyl) ester			
2,3-Pentanedione			
Benzene, 3,5-dimethyl-1-(phenylmethyl)-			

The end of Table 2

Control group	Gastric carcinoma	Lung cancer	Colon cancer
1-(1-Trimethylsilyloxyethenyl)-3-trimethylsilyloxy-benzene			
3,5-Dihydroxybenzamide			
15-Crown-5			

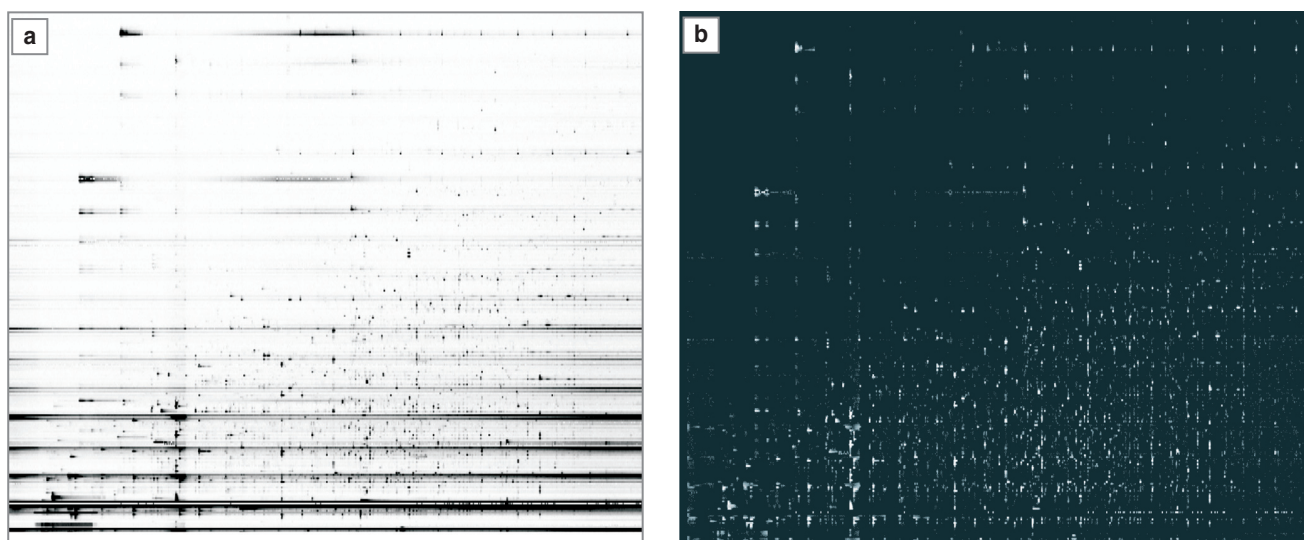


Fig. 3. An example of interferences eliminated from a mass chromatogram image: *a* — original image; *b* — an image after interference elimination

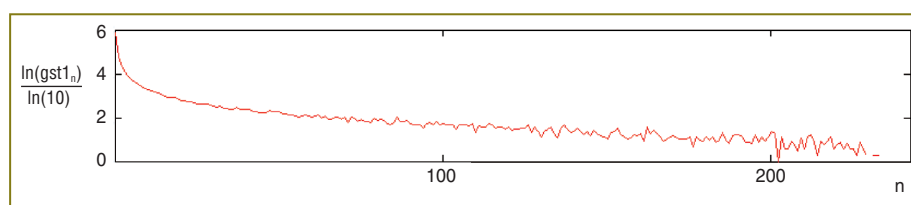


Fig. 4. An example of an intensity histogram of a processed image of a mass chromatogram. Along Y direction logarithmic scale is used

Reference mass-spectra are taken from mass-spectral data libraries NIST 02, NIST 05, WILEY, etc., which are connected to the programs generally being a part of mass-spectrometric equipment.

Then we determined the detection rate of each metabolite in different groups and formed their metabolic profiles. We chosen the metabolites, the number of which was 20% more in one group compared to any other group. After the procedure metabolic profiles for each disease and a control group were revealed as shown in Table 2.

The decoding of chromat-mass-spectrometric data can be accompanied by errors due to ill-defined chromatographic peaks, errors of identification of substances according to mass-spectra, incompleteness

of the libraries, etc. Therefore, we considered a possibility to form metabolic profiles of cancer diseases according to the images of mass chromatograms of urine volatile organic substances.

The procedure of forming the images of mass chromatograms corresponding to different oncological diseases consists

of several stages. Except relevant data an initial image of a mass chromatogram also contains a large proportion of interferences due to the fact that some materials from the column inner layer enter the detector, as well as due to the impurities the carrier gas has, etc. On an image they are evident as horizontal lines throughout a mass chromatogram (See Fig. 2). Moreover, there can be mass chromatograph base-line drift that can be shown as the changes of mean brightness in the course of time. For interference elimination at the first stage an image was filtrated by a horizontal strobe, which eliminated the steady component along the image line. Brightness values of background counts in the resulting image were determined by the following formula suggested by the authors:

$$J_{x,y}^{proc} = \begin{cases} \sum_{n=-s/2}^{s/2} J_{x+n,y} & , \text{ if the result is above zero;} \\ 0 & \text{ otherwise.} \end{cases}$$

where $J_{x,y}$ — pixel brightness of an original image of a spectrogram with coordinates x, y ; s — filter aperture size; $J_{x,y}^{proc}$ — pixel brightness of a processed image of a spectrogram with coordinates x, y .

Fig. 3 shows an example of mass chromatogram image processing.

The next step is the thresholding of the obtained image in order to reveal the most characteristic metabolites of a patient. According to the analysis of intensity histogram (Fig. 4), the most part of background counts in a mass



Fig. 5. An example of threshold processing of a mass spectrogram image

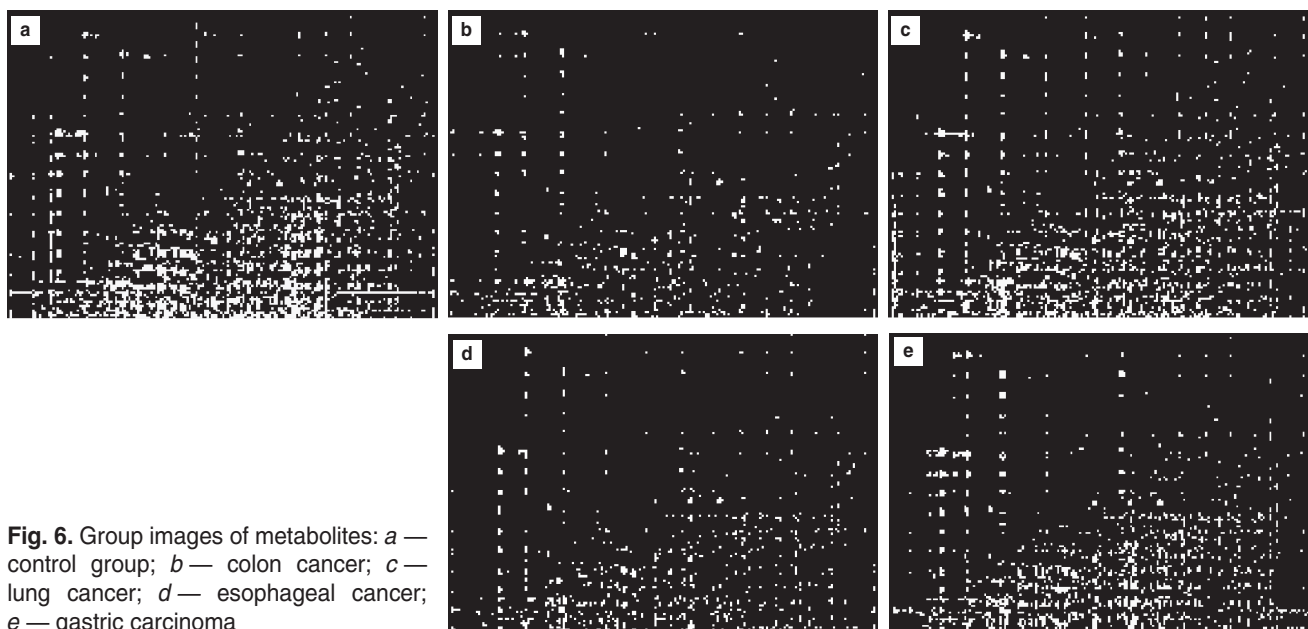


Fig. 6. Group images of metabolites: a — control group; b — colon cancer; c — lung cancer; d — esophageal cancer; e — gastric carcinoma

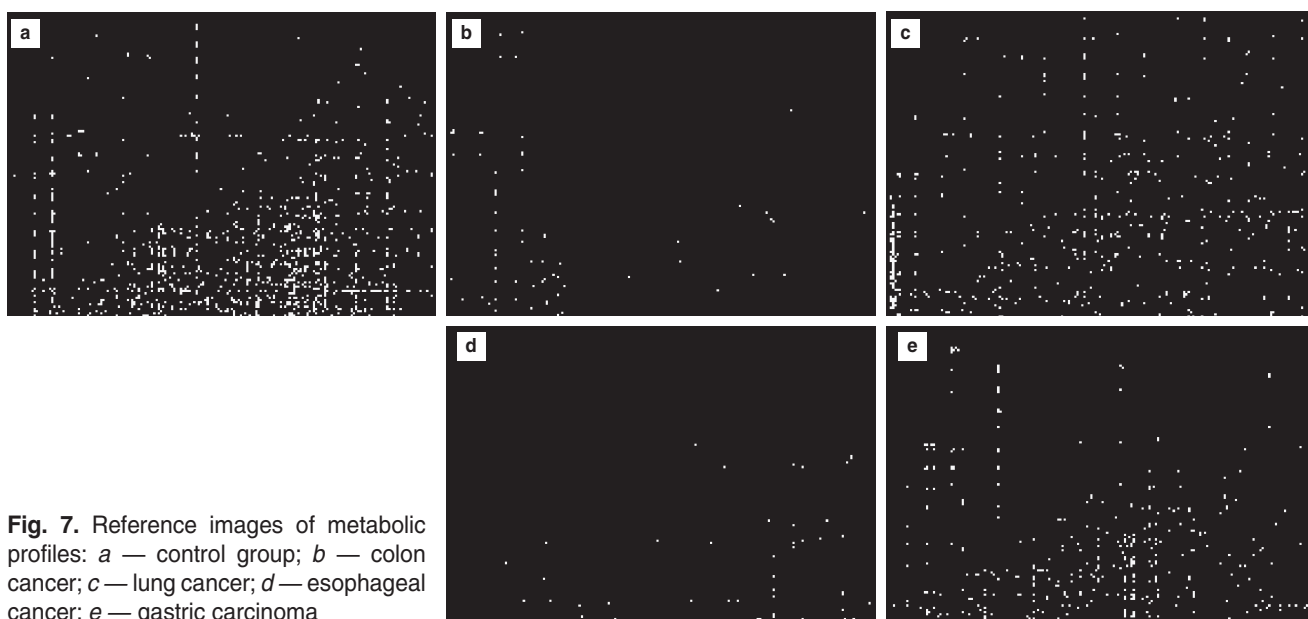


Fig. 7. Reference images of metabolic profiles: a — control group; b — colon cancer; c — lung cancer; d — esophageal cancer; e — gastric carcinoma

chromatogram image is in the intensity range from 0 to 50, therefore, the threshold of brightness cutoff was chosen equal to 50. Fig. 5 shows an example of threshold processing of an image represented in Fig. 3, b.

For metabolic profile formation of a certain cancer type disease according to the present technique, we processed all images of mass chromatograms of patients of a certain group, as well as of the control group. After that we formed group images containing information about all metabolites the patients of the group had. For this purpose we performed the operation of logical OR between all images of this group. Fig. 6 shows the examples of such images for the control group, patients with lung cancer, esophageal cancer, gastric carcinoma, and colon cancer.

The next processing stage consisted in revealing unique elements and formation of reference images of metabolic profiles according to the rule:

$$J_{x,y}^{ref_m} = \begin{cases} 1, & \text{if } J_{x,y}^{proc_n} = 0, m \neq n \text{ and } J_{x,y}^{proc_m} \neq 0, n=1, \dots, 5; m=1, \dots, 5; \\ 0 & \text{otherwise.} \end{cases}$$

where $J_{x,y}^{ref_m}$ — pixel brightness in a reference metabolic profile of m disease, $J_{x,y}^{proc_n}$ — pixel brightness of a processed image of a spectrogram of n disease with coordinates x, y . Thus, a resulting image of this group of patients contains only those metabolites, the images of other groups have not. Fig. 7 demonstrates the images of mass chromatograms of reference metabolic profiles.

Results. For capability assessment of identification of different cancer types according to the analysis of urine volatile organic substances we carried out an experiment. Using statistical methods we analyzed the efficacy of identification of different cancer forms. For this purpose we compared a mass chromatogram image of each patient with the images of reference mass chromatograms (See Fig. 7). The reference, which had the highest similarity, was taken as the result of identification. After processing of the sampling we assessed the probability of correct and wrong solutions (Table 3).

According to the assessment of the obtained results, the technique sensitivity of the analyzed sampling was 100%, specificity — 90.62%, type I error probability — 0, type II error probability — 9.38%.

Discussion. Currently, the well known and most common approaches to the problem of early diagnosis of cancer diseases are the following screening methods:

1. Radiological methods include traditional radiological investigations, X-ray tomography, ultrasound investigation, tomography (computed, magnetic resonance, positron emission tomography), radioisotopic (radionuclide) studies. Diagnostic value of these techniques is high; however, their application requires highly qualified staff, the use of expensive equipment with limited capacity. Therefore, to apply these methods for population screening studies is not always reasonable.

2. Endoscopic examination. It is generally used to detect cancer diseases affecting mucosa of the organs, which are accessible per vias naturals, and this limits their application.

3. Metabolic profiling techniques are based on the human metabolic changes in cancer. Various methods are used for metabolite composition analyses, including magnetic resonance spectroscopy (MRS), nuclear magnetic resonance — NMR (mainly, ^1H NMR), Fourier ion cyclotron resonance mass-spectrometry, as well as other mass-spectrometric techniques, which are frequently combined with separation methods — liquid chromatography and gas chromatography. MRS can be used supplementary to MRT and enables to detect specific metabolites under natural conditions. The main MRS advantage is its capability to reveal and measure tissue-specific metabolites, their drawbacks are low sensitivity and low spectral resolution.

Among all the methods mentioned, mass-spectrometry is the most sensitive technique that is why its application in clinical practice is prospective.

Except for the selection of optimal methods, one more important component of cancer diagnosis is the material used. In this regard, the following techniques are of interest.

1. Expired air composition analysis. Most studies in this field belong to the team lead by M. Phillips [7, 8]. They developed “a breath analyzer” for automated breath analysis and detection of volatile organic compounds for different cancer types. The best results were obtained in lung cancer diagnosis (84.5% — sensitivity, and 81.0% — specificity), however, there is no data on other cancer types prompting suggestions that in these

Table 3
Diagnostic efficiency assessment

Groups	Control group (n=35)	Colon cancer (n=8)	Lung cancer (n=14)	Esophageal cancer (n=12)	Gastric carcinoma (n=12)	Right diagnosis, %
Control group	29	0	3	0	0	90.625
Colon cancer	0	9	2	1	0	75
Lung cancer	0	0	28	1	0	96.55172414
Esophageal cancer	0	0	0	8	0	100
Gastric carcinoma	0	0	0	1	26	96.2962963

cases the technique is non-effective. Thus, there is the necessity of developing effective standard techniques for sample preparation.

2. Tissue composition analysis. The technique is based on the change of chemical tissue composition in case of pathology. It can be applied to diagnose esophageal cancer, gastric carcinoma, ovarian carcinoma, malignant and benign breast tumors. The drawback of the method is invasiveness of sample preparation procedures, the difficulty of sample preparation before analysis.

3. Blood count. In contrast to tissue composition analysis, blood count has many advantages in sample acquisition. On the other hand, samples are sensitive to the change of temperature, pH, ion concentration, etc. The disadvantage of the approach is the difficulty of sample preparation, and significant dependence of the results obtained on experimental conditions. There is no available data on possible identification of a wide range of cancer diseases.

Enzyme immunoassay of blood lacks the mentioned drawbacks. However, a complex study of several cancer diseases is rather expensive.

4. Urinalysis. There are researches on using gas chromatography-mass spectrometry to detect bladder, breast, ovarian, cervical cancer, renal carcinoma, hepatocellular carcinoma. Most studies concern the identification of some cancer types, good diagnostic results have been obtained for some cancer types. However, these studies do not consider the problem of identification of one cancer types among others, and the efficacy of the suggested techniques for these cases has not been analyzed.

The technique we proposed is based on the determination of cancer-specific markers according to the composition of volatile organic compounds of urine, the technique is free of the advantages mentioned. The proposed technique is easy, non-invasive, low labor intensive and affordable; it makes it applicable for population screening programs.

Conclusion. The suggested technique of cancer detection based on image analysis of mass chromatograms of volatile organic compounds of urine is high-efficient. The technique sensitivity is 100%, specificity — 90.62%.

The present findings can serve the base for researches in the diagnosis of different cancer types, and can be used in clinical practice for primary cancer detection.

Study Funding and Conflict of Interests. The study was not supported by any funds, and the authors have no conflict of interest to disclose.

Acknowledgment. The authors express their thanks to Chromatec for equipment provided and research assistance.

References

1. *Statistika zaboлеваemosti rakom [Cancer statistics]*. <http://www.knigamedika.ru/novooobrazovaniya-onkologiya/statistika-zaboлеваemosti-rakom.html>.
2. Furina R.R., Mitrakova N.N., Ryzhkov V.L., Safiullin I.K. Metabolomic research in medicine. *Kazanskij medicinskij zurnal* 2014; XCV(1): 1–6.
3. Oliveira P.A., Colaco A., Chaves H.R., et al. Chemical carcinogenesis. *An Acad Bras Cienc* 2007; 79(4): 593–616, <http://dx.doi.org/10.1590/S0001-37652007000400004>.
4. Tsarev N.I., Tsarev V.I., Katrakov I.B. *Prakticheskaya gazovaya khromatografiya [Practical Gas Chromatography]*. Barnaul: Izd-vo Alt. un-ta; 2000; 156 p.
5. Patti G.J., Yanes O., Siuzdak G. Innovation: metabolomics: the apogee of the omics trilogy. *Nat Rev Mol Cell Biol* 2012; 13(4): 263–269, <http://dx.doi.org/10.1038/nrm3314>.
6. Zimmermann D., Hartmann M., Moyer M.P., et al. Determination of volatile products of human colon cell line metabolism by GC/MS analysis. *Metabolomics* 2007; 3(1): 13–17, <http://dx.doi.org/10.1007/s11306-006-0038-y>.
7. Phillips M., Cataneo R.N., Cummin A.R., et al. Detection of lung cancer with volatile markers in the breath. *Chest* 2003; 123(6): 2115–2123, <http://dx.doi.org/10.1378/chest.123.6.2115>.
8. Phillips M., Gleeson K., Hughes J.M., et al. Volatile organic compounds in breath as markers of lung cancer: a cross-sectional study. *Lancet* 1999; 353: 1930–1933, [http://dx.doi.org/10.1016/S0140-6736\(98\)07552-7](http://dx.doi.org/10.1016/S0140-6736(98)07552-7).
9. Woo H.M., Kim K.M., Choi M.H., et al. Mass spectrometry based metabolomic approaches in urinary biomarker study of women's cancers. *Clin Chim Acta* 2009; 400(1–2): 63–69.
10. Kouremenos K.A., Pitt J., Marriott P.J. Metabolic profiling of infant urine using comprehensive two-dimensional gas chromatography: Application to the diagnosis of organic acidurias and biomarker discovery. *J Chromatogr A* 2010; 1217(1): 104–111, <http://dx.doi.org/10.1016/j.chroma.2009.10.033>.
11. Pauling L., Robinson A.B., Teranishi R., Cary P. Quantitative analysis of urine vapor and breath by gas-liquid partition chromatography. *Proc Natl Acad Sci USA* 1971; 68: 2374–2376.
12. Silva C.L., Passos M., Câmara J.S. Investigation of urinary volatile organic metabolites as potential cancer biomarkers by solid-phase microextraction in combination with gas chromatography-mass spectrometry. *Br J Cancer* 2011; 105(12): 1894–1904, <http://dx.doi.org/10.1038/bjc.2011.437>.