

Processing and Interpretation of Analytical Data with a High Degree of Uncertainty

I. G. Zenkevich^{a, *}, D. A. Nikitina^a, and A. S. Kushakova^b

^a Institute of Chemistry, St. Petersburg State University, St. Petersburg, 198504 Russia

^b ZAO Biokad, St. Petersburg, 198515 Russia

* e-mail: izenkevich@yandex.ru

Received December 8, 2021; revised January 8, 2022; accepted March 28, 2022

Abstract—Some analytical problems, which are often considered incorrect for objective reasons, are considered. The main of these reasons is an anomalously large scatter of the initial data. It can be due to either the low reproducibility of the characteristics of substances, their quantities, analytical signal intensities, process conditions, etc., or variability due to differences in the nature of the objects themselves. In the latter case, the nature of data interpretation is influenced by analytical hypotheses adopted for their consideration. The tasks considered include variations in the component composition of developers for black-and-white negative photographic materials, comparison of temperature parameters of the gas-chromatographic separation of various organic compounds, toxicity characterization (LD₅₀) of homologues using C₃–C₁₂ 1-alkanols as an example, and the possibility of predicting sample preparation operations in the determination of drugs in blood plasma based on their physicochemical characteristics. The main features of data interpretation characterized by a high degree of uncertainty are revealed. It is noted that important conclusions can be drawn based on the facts of low reproducibility (one-dimensional arrays) or poor correlation of variables (two-dimensional arrays).

Keywords: data arrays, high uncertainty, interpretation, examples of various tasks, drug preparations, blood plasma, sample preparation, hydrophobicity factors

DOI: 10.1134/S1061934822090143

The most important requirements for analytical methods (a frequently used term is *validation*), traditionally include such a characteristic as the precision of measurement results subdivided into their repeatability and reproducibility in addition to accuracy, robustness, etc. The evaluation criterion is the minimum standard deviations of the obtained values (or coefficients of variation). Two-dimensional data arrays (first of all, calibration dependences) are characterized by linearity, the criterion of which is the maximum values of correlation coefficients (R) or the minimum values of general variances (S_0) [1]. Many practice guidelines specify particular accuracy performance requirements. This is often referred to bioanalytical methods and drug control [2–4]. For example, in accordance with rules [4], the permissible deviations of the results for at least 67% samples should be within $\pm 15\%$ of their nominal values; otherwise, the reasons for the deviations should be clarified.

However, in real analytical practice, there are a large number of examples of data arrays that are characterized by low reproducibility or poor correlation of variables (in other words, a high degree of uncertainty). This may be due to either a large scatter of direct measurement results or variability because of differences in the nature of objects. Spark source mass

spectrometry [5, 6] is an example of an objectively large spread of signal intensities; the same features are often inherent in determinations at a trace level. The general approach to eliminating such uncertainties is to improve the technique of determinations. Variability of the second type sometimes is fundamentally irremovable, but it depends on the accepted formulations of analytical problems (in other words, on accepted analytical hypotheses). As an example of the data scatter caused by the data sample, we can cite the intensities of weak signals of so-called mass spectra of ion series [7]¹. Their standard deviations may exceed the mean values.

Another example of a weak correlation of values is the assessment of retention indices (RIs) in reversed phase HPLC based on the hydrophobicity factors ($\log P$) of analytes [8, 9]. As a rule, the correlation coefficients of the RI– $\log P$ relationships are at a level of 0.9, and S_0 (the average accuracy of the obtained estimates) is at least 50–70 index units. Nevertheless, this approach has been used to date due to the relative availability of (predominantly calculated) values of $\log P$. Such correlations are often regarded as having

¹ The spectra of ion series combine data for individual homologues into averaged characteristics of homologous series [7].

only limited application or even as unsatisfactory. However, the numerous examples of a large scatter of data not only do not allow them to be neglected but also deserve special consideration of the interpretation of such information, which is the goal of this work.

To illustrate the diversity of such examples, we analyzed in detail four different problems. All of them refer to variability due to differences in the nature (properties, characteristics) of objects, and a strong dependence of the features of data interpretation on the formulation of problems under consideration is manifested in all of them. As the simplest example (one-dimensional data array), variations in the composition of developers for black-and-white negative photographic materials are considered. More complex problems include variations in the temperature regimes of gas-chromatographic analysis because the problem becomes two-dimensional when data for different analytes are combined. The toxicity of homologues is another example of large scatter. Finally, the analytically important problem of choosing sample preparation operations for determining drug traces in blood plasma based on the physicochemical characteristics of the test compounds is considered in more detail.

CHOICE OF INITIAL DATA AND SAMPLE PREPARATION FOR ANALYSIS (EXPERIMENTAL)

General characteristics of the four considered examples. To exclude the search for and comparison of information from insufficiently systematized sources of different times of publication, information on the composition of developers for black-and-white negative photographic materials (example 1) was taken from a manual [10], which was published in a period of a great variety of both negative photographic materials themselves and recipes of processing solutions for them (1964). The database of the National Institute of Standards and Technology (NIST, the United States) (<http://webbook.nist.gov>) [11] was chosen as a source of information on the conditions of gas-chromatographic analysis with temperature programming for five analytes (nitrobenzene, benzonitrile, 2-chlorophenol, 1,4-dimethoxybenzene, and 1,3,5-trichlorobenzene); this database contains detailed information on all compounds whose retention indices are included in it (example 2). The selected parameters include the initial temperature (T_0 , °C), the duration of an initial isothermal section ($t(T_0)$, min), and the heating rate (r , K/min). Toxicity data (the values of LD_{50} , example 3) for C_3 – C_{12} 1-alkanols were taken from various literature sources [12–18]. Many of them contain duplicate values of LD_{50} , which were not detected and rejected. Finally, the descriptions of sample preparation procedures for drugs (51 compounds of various chemical nature) are summarized on the basis of original publications and data obtained

from JSC Biokad (see below) (example 4). Tables 1–4, the numbers of which are consistent with the numbers of examples, summarize experimental data corresponding to each of the examples and necessary references. The values of hydrophobicity factors ($\log P$) and the degrees of drug binding to proteins are indicated on the Drugbank website [62].

Preparation of samples for analysis. Data in Table 4, which corresponds to the last considered example and contains information on the operations of blood plasma sample preparation for HPLC determination of drugs, is supplemented with information for several drugs the analysis methods for which were developed at ZAO Biokad (St. Petersburg). Their standard solutions with a concentration of 1 mg/mL were prepared by dissolving weighed portions of the analyzed substances in a mobile phase (the composition of the solvents corresponded to the initial compositions of eluants in gradient elution modes). The blood plasma of healthy volunteers stored frozen at a temperature of no higher than -70°C was used as a matrix for preparing model solutions.

Samples for HPLC analysis were prepared according to one of the following procedures: (1) Liquid–liquid extraction (LLE) was carried out by adding an extractant to aliquot portions of the samples (1 : 3); the resulting solutions were mixed and centrifuged; the organic extracts were evaporated in a flow of nitrogen, and the dry residues were repeatedly dissolved in the mobile phase. (2) For the precipitation of blood plasma proteins, acetonitrile was used as a precipitant in a ratio of 1 : 3; the resulting solutions were mixed and centrifuged, and supernatant layers were transferred into vials for chromatographic analysis. (3) Solid phase extraction (SPE) was performed using cartridges (Oasis, Waters, the United States); samples were applied to them and various solvents were passed through the cartridges to elute impurities and target analytes separately. (4) Amicon® Ultra 3K centrifuge ultrafilters (Millipore, the United States) were used to filter plasma proteins. Specified volumes of the samples were added to test tubes with filters and centrifuged; the amounts of solutions passed through the membranes were taken for chromatographic analysis.

In all the above examples, the initial information was highly variable, and the scatter of numerical data was large so that any correction of single values in such arrays (identification of outliers) [63] was impossible. The ORIGIN software (versions 4.1 and 8.1) was used for statistical processing of the initial data and plotting. The significant variability of the initial data limits the use of more complex methods of their interpretation (factor or cluster analysis). Moreover, in all of the examples, the number of available values is insufficient for the correct application of these methods.

Table 1. Concentrations of components in metol–hydroquinone developers for negative black-and-white photographic materials (according to Yashtold-Govorko [10])

Developer	Component concentration, g/L					
	metol	hydroquinone	sodium sulfite (crystal)	sodium carbonate (anhydrous)*	potassium bromide	other additives
NP-1	2.5	2.5	60	5	0.5	–
Supporting solution for NP-1	7	10	60	20	–	–
NP-2	1	5	52	20	1	–
NP-3	5	6	80	40	3	–
Contrast	–	30	120	–	20	19 (NaOH)
Contrast KT-1	2	10	104	40	4	–
For high temperatures	6	–	180	–	2	23 (borax), 100 (Na ₂ SO ₄)
Fine-grained NP-15	2	5	200	–	–	2 (borax)
Supporting solution for NP-15	3	7.5	200	–	–	20 (borax)
Standard	8	–	250	5.75	2.5	–
NP-19A	10	–	120	90	2	–
Fine-grained D-23	7.5	–	200	–	–	–
DK-20	5	–	200	–	0.5	2 (borax), 1 (KNCS)
For equalizing image contrast	5	6	40	25	4.5	–
For correcting underexposures	14	14	50	–	9	9 (NaOH)
For increasing light sensitivity	1	5	52	20	1	0.2 (hydrazine sulfate), 0.01 (benzotriazole)
For overexposed photographic materials	7	–	50	12	5	–
For fast development	5	45	180	–	10	40 (NaOH), 1 (benzotriazole)
Average values $\pm s$	5 ± 4	$8 \pm 12^{**}$	120 ± 70	$15 \pm 23^{**}$	$4 \pm 5^{**}$	–
Sample asymmetry (<i>A</i>)	1.6	3.5	1.5	2.3	2.9	–

* In more recent developer formulations, hygroscopic sodium carbonate was replaced by nonhygroscopic potassium carbonate (conversion factor, 1.30).

** Components for which standard deviations exceed the arithmetic mean values of their concentrations in solution.

RESULTS AND DISCUSSION

Component composition of metol–hydroquinone developers for negative black-and-white photographic materials. It is advisable to start considering data sets with a large scatter with one-dimensional arrays to which conventional statistical processing is applicable. According to modern ideas, the best results of processing photographic materials are provided by strict reproduction of the conditions, including the composition of solutions, the duration of processing, and the temperature regime (maximal standardization). However, the situation was fundamentally different 50–60 years ago: for numerous negative photographic mate-

rials (with different spectral sensitivity), a large number of developers were used for various purposes, including normal (standard), fine-grained, contrast developers and those to reduce contrast, to increase light sensitivity, for overexposed photographic materials, for high processing temperatures, etc. All of these developers differed in combinations and concentrations of components. If we confine ourselves only to metol–hydroquinone developers² for negative photographic materials, Table 1 summarizes the concentra-

² Metol, 4-(methylamino)phenol sulfate; hydroquinone, 1,4-dihydroxybenzene.

Table 2. Comparison of the parameters of temperature regimes for gas-chromatographic analysis of some organic compounds (according to NIST Standard Reference Database [11])

Compound	T_0 , °C	$t(T_0)$, min	r , K/min	Compound	T_0 , °C	$t(T_0)$, min	r , K/min
Nitrobenzene	40	0	5	1,4-Dimethoxybenzene	60	2	10
	60	1	5		60	5	3
	75	0	6		35	3	4
	50	2	8		60	1	6
	40	1	10		40	7	6
Average values	55 ± 13	1 ± 1	7 ± 2		60	0	3
Benzonitrile	0	12	12		40	5	3
	40	0	5		70	2	3
	40	0	3		35	4	2
	50	0	2.5		60	0	4
	50	0	6		50	0	4
	35	2	2		70	5	4
Average values	36 ± 19	2 ± 5	5 ± 4		80	0	3
2-Chlorophenol	40	0	3		50	2	3
	35	0	8		40	10	3
	50	0	10	Average values	54 ± 14	3 ± 3	4 ± 2
	50	0	5	1,3,5-Trichlorobenzene	50	0	2
	50	4	6		50	0	2
	100	0	10		100	0	10
	100	0	2		100	0	5
	100	0	6		50	0	10
	100	0	10		50	0	5
	100	0	5		30	0	1
	50	0	10		80	0	2
	50	0	5		80	0	3
	50	3	3		40	3	2
	80	0	2	80	0	2	
	50	2	8	50	5	5	
	40	3	8	Average values	63 ± 23	1 ± 2	4 ± 3
	50	4	6				
Average values	64 ± 25	1 ± 2	6 ± 3				

Overall mean values: $T_0 = (57 \pm 22)^\circ\text{C}$ ($A = 1.6$), $t(T_0) = (2 \pm 3)$ min ($A = 2.5$), and $r = (5 \pm 3)$ K/min ($A = 1.7$).

tions of components in 19 of them according to a manual [10].

The purpose of such a variety of compositions is to choose the best option for processing depending on the characteristics of a particular photographic material or shooting conditions. However, another real situation can be imagined: the type of photographic material, the features of its exposure, and, consequently, the a priori requirements for processing are unknown. In other words, here we are dealing with a different formulation of the original hypothesis. In such cases, the choice of a specific developer formula-

tion from among the previously known ones is clearly irrational; therefore, a change in the logical scheme of actions is required. From the uncertainty of the formulation of the problem follows the preferred use of a formulation the composition of which corresponds to the average amounts of all components (arithmetic averages are given in the last row of Table 1).

Variations in the concentrations of individual components (ranges) are very large: 0–14 g/L (metol), 0–45 g/L (hydroquinone), 40–250 g/L (sodium sulfite), 0–90 g/L (sodium carbonate), and 0–20 g/L (potassium bromide). Not surprisingly, the formal calcula-

Table 3. Data on the toxicity of C₃–C₁₂ 1-alkanols (LD₅₀, mg/kg, p.o.; no indication of the animal corresponds to the use of rats)

1-Alkanol	LD ₅₀	Reference
1-Propanol	1870 2160, 2800 (rabbit), 1870 1870, 5700 (for humans) 4400, 7000 2825 (rabbit), 6800 (mouse), 1870, 2200	SDS* (Thermo Fisher Scientific, 2018) PubChem SDS (ChemSupply, Australia) SDS (Thermo Fisher Scientific) SDS (ASS no. 01356)
Average value (<i>A</i>)	3400 ± 2000 (2.2)	
1-Butanol	2680, 2510, 2020, 790, 1200 (hamster), 2500 (bird), 3500 (rabbit), 1782 (dog) 4360, 2290, 2510, 2020, 790, 2680 (mouse), 3500 (rabbit), 1200 (hamster), 1782 (dog) 700, 790 2292, 3430 (rabbit) 790–4360 2500, 3400 (rabbit), 790, 2680 (mouse), 3484, 2500 (bird) 1200 (hamster), 2680 (mouse), 3400 (rabbit), 3500 (rabbit), 2100, 800–2000, 700 2292 (rat) 2290, 2510, 2680 (mouse), 3500 (rabbit), 1200 (hamster), 1782 (dog), <2000	[12] [13] SDS (Thermo Fisher Scientific) SDS (Carl Roth, 2021) PubChem SDS (Sigma-Aldrich) [14] [15] [16]
Average value (<i>A</i>)	2200 ± 1000 (0.90)	
1-Pentanol	200 (mouse). 370 >2000 >2000 3645 3670, 200 (mouse) >2000 2200–3000 2200, 3600, 140–4585	SDS (Thermo Fisher Scientific) SDS (Carl Roth) SDS (Acros Organics) SDS (Agilent) SDS (Sigma-Aldrich) SDS (ChemSupply, Australia) SDS (Calderon Lab Chemicals) [17]
Average value (<i>A</i>)	2100 ± 1500 (0.97)	
1-Hexanol	1950 (mouse) 720, >2000 4590, 4870, 103–4870 200–2000 500–5000 >2000	[12] SDS (Carl Roth) [17] [18] SDS (Cameo Chemicals) ChemBook
Average value (<i>A</i>)	2400 ± 1900 (1.8)	
1-Heptanol	3250, 6200, 5500, 500–6200 4300 (mouse) 1870 500 (rabbit), >2000 500	[17] PubChem SDS (Cameo Chemicals) ChemBook Green Chemistry
1-Octanol	>3200 >5000, 1790 →5000 >2000	[12] [11] ChemBook
1-Nonanol	3560, 800–6400	[17]
1-Decanol	4720, 1000–5000	[17]
1-Undecanol	3000, 3000→15800**	[17]
1-Dodecanol	>2000, 1500–26530**	[17]

* SDS (Safety Data Sheet) or MSDS (Material Safety Data Sheet) is the designation of websites of various companies containing the toxicological characteristics of chemical compounds.

** Abnormally large values of LD₅₀ excluded from the graphical representation of the data (Fig. 2).

Table 4. Main characteristics of sample preparation operations for drug preparations depending on their hydrophobicity factors ($\log P$) (compounds are arranged in a descending order of the calculated values of $\log P$ (ChemAxon))

Name	Molecular weight	CAS no.	$\log P$	Binding with plasma proteins, %	Sample preparation technique*	References
Everolimus	958	159351-69-6	7.4 (ChemAxon)	~74%	LLE**	
Eltrombopag	443	496775-61-2	6.03 (ChemAxon)	~99%	Precipitation	[19]
Atorvastatin	559	134523-00-5	5.39 (ChemAxon) 6.36 (expt)	~98%	LLE	[20]
Bosentan	552	147536-97-8; 157212-55-0 (monohydrate)	4.94 (ChemAxon) 3.7 (expt)	~98%	Precipitation**	
Dapoxetine	305	119356-77-3	4.67 (ChemAxon)	~99%	LLE	[21]
Vinorelbine	779	71486-22-1	4.65 (ChemAxon) 4.0 (expt)	~90%	LLE	[22]
Simeprevir	750	923604-59-5	4.56 (ChemAxon)	~99%	LLE	[23]
Loratadine	383	79794-75-5	4.55 ChemAxon	99%	LLE	[24]
Pyronaridine	518	74847-35-1	4.22 (ChemAxon)	~92%	LLE	[25]
Nateglinide	317	105816-04-4	4.03 ChemAxon	~98%	Micro-LLE	[26]
Abiraterone	392	154229-18-2	3.97 (ChemAxon)	~99%	LLE**	
Chloroquine	320	54-05-7	3.93 (ChemAxon) 4.63 (expt)	~74%	Precipitation	[25]
Gefitinib	447	184475-35-2	3.75 (ChemAxon) 3.2 (expt)	~90%	LLE**	
Crizotinib	450	877399-52-5	3.57 (ChemAxon)	~91%	Precipitation	[27, 28]
Pazopanib	438	444731-52-6 635702-64-6 (hydrochloride)	3.55 (ChemAxon) 3.2 (expt)	~99%	Precipitation**	
Mycophenolic acid	320	24280-93-1	3.53 (ChemAxon) 2.8 (expt)	~98%	LLE	[29]
Canagliflozin	444	842133-18-0	3.52 (ChemAxon)	~99%	LLE	[30]
Artemether	298	71963-77-4	3.48 (ChemAxon)	~95%	Precipitation	[25]
Olanzapine	312	132539-06-1	3.39 (ChemAxon)	~93%	LLE	[31]
Asunaprevir	748	630420-16-5	3.37 (ChemAxon)	~99%	LLE	[32]
Bupropion	276	34911-55-2	3.27 (ChemAxon) 3.6 (expt)	~85%	LLE	[33]
Erlotinib	393	183321-74-6	3.2 (ChemAxon)	~95%	Precipitation**	
Stiripentol	234	49763-96-4	3.12 (ChemAxon)	~99%	LLE Precipitation	[34] [35]
Diazepam	285	439-14-5	3.08 (ChemAxon)	~98%	LLE	[36]
Sunitinib	398.5	557795-19-4	2.93 (ChemAxon)	~95%	LLE	[37]
Docetaxel	808	114977-28-5	2.92 (ChemAxon) 2.4 (expt)	~97%	LLE	[38]
Fluvoxamine	318	54739-18-3	2.8 (ChemAxon) 3.2 (expt)	~80%	LLE	[39]

Table 4. (Contd.)

Name	Molecular weight	CAS no.	log <i>P</i>	Binding with plasma proteins, %	Sample preparation technique*	References
Gilteritinib	553	1254053-43-4	2.79 (ChemAxon)	~94%	SPE	[40]
Retigabine	303	150812-12-7	2.7 (ChemAxon)	~80%	LLE	[41]
Eslicarbazepine acetate	296	236395-14-5	2.17 (ChemAxon)	~40%	Precipitation + salting-out	[42]
Teriflunomide	270	163451-81-8	2.14 (ChemAxon) 1.536 (expt)	~99%	Precipitation**	
Dapagliflozin	409	461432-26-8	2.11 (ChemAxon) 2.7 (expt)	~91%	LLE	[30]
Oxcarbazepine	252	28721-07-5	1.82 (ChemAxon)	~40%	LLE	[43]
Empagliflozin	451	864070-44-0	1.66 (ChemAxon)	~86%	LLE	[30]
Dolutegravir	419	1051375-16-6	1.1 (ChemAxon)	~99%	Precipitation + SPE	[44]
Doxorubicin	544	23214-92-8	0.92 (ChemAxon)	~74%	Precipitation	[45]
Monomethyl fumarate	130	2756-87-8	0.34 (ChemAxon)	~40%	Precipitation**	
Zonisamide	212	68291-97-4	0.11 (ChemAxon) 0.5 (expt)	~40%	LLE	[46, 47]
Cisplatin***	300	15663-27-1	-2.19 (Hansch, 1995) -0.04 (ALOGPS)	~95%	Derivatization + LLE	[48]
Mercaptopurine	152	50-44-2	-0.12 (ChemAxon)	~19%	Precipitation	[49]
Sulthiame	290	61-56-3	-0.27 (ChemAxon)	~60%	Precipitation	[50]
Didanosine	236	69655-05-6	-0.35 (ChemAxon)	~5%	SPE	[51]
Metronidazole	171	443-48-1	-0.46 (ChemAxon)	~20%	Precipitation	[52]
Isoniazid	137	54-85-3	-0.69 (ChemAxon)	~10%	Precipitation	[53]
Lenalidomide	259	191732-72-6	-0.71 (ChemAxon) -0.4 (expt)	~30%	Precipitation**	
Entecavir	277	142217-69-4 209216-23-9 (monohydrate)	-1.4 (ChemAxon) -0.8 (expt)	~13%	Ultrafiltration**	
Cytarabine	243	147-94-4	-2.8 (ChemAxon) -2.8 (expt)	~13%	Cation exchange SPE	[54]
Azacitidine	244	320-67-2	-3.1 (ChemAxon) -3.5 (expt)	N/A	Cation-exchange SPE	[55]
					Precipitation	[56]
Tigecycline	585.6	220620-09-7	-3.9 (ChemAxon) 0.8 (expt)	~89%	Precipitation	[57]
Ceftazidime	547	72558-82-8	-4.1 (ChemAxon)	~10%	Ultrafiltration	[58]
Meropenem	384	96036-03-2	-4.4 (ChemAxon) -0.6 (expt)	~2%	Precipitation	[59, 60]
					SPE	[61]

* Abbreviations for sample preparation operations: LLE, liquid-liquid extraction, SPE, solid-phase extraction.

** References are not indicated for preparations characterized at ZAO Biokad.

*** Inorganic complex compound; experimental values of log *P* are unknown.

tion of standard deviations leads to values that exceed their arithmetic mean values for some of them (hydroquinone, sodium carbonate, and potassium bromide). For the other two components, the relative standard deviations (coefficients of variation) are 80% (metol) and 58% (sodium sulfite). In the practice of mathematical processing of measurement results [64], if the standard deviation of a value is greater than its mean value, this means that such a value is statistically insignificant, and it can be assumed equal to zero. However, in relation to the problem under consideration, such an interpretation is unacceptable, and it should be changed: this confirms the possibility of the absence of such a component from the developer in some cases.

If the considered data arrays $\{x_j\}$ are obviously asymmetric (as in this case), they can be characterized by an asymmetry factor (A). One way to estimate this parameter is to calculate two standard deviations independently. The first characterizes data that exceed the arithmetic mean, $s(+)$, and the second, accordingly, smaller, $s(-)$. Their ratio $A = s(+)/s(-)$ is a characteristic of asymmetry [65]. In our case, all values of A (indicated in the last row of Table 1) are significantly greater than unity (from 1.5 to 3.5). This is typical for all data arrays bounded from below by zero, while there are no such restrictions from above.

Such a scatter of data on the composition of developers deserves comment. First of all, it corresponds to an average recipe. If we consider the 11 specific compositions listed in Table 1, the calculated average values correspond best to developers NP-3 and for equalizing image contrast. In addition, note that if we estimate the accuracy of setting the required quantities of ingredients by eye approximately as ± 50 rel %, the standard deviations of the obtained mean values exceed this value. This means that it is possible to exclude an operation of taking accurate weights of components from the preparation of these developers; nevertheless, this ensures an acceptable level of quality in the processing of photographic materials.

As noted above, similar examples of the statistical processing of data characterized by a noticeable scatter occur in the calculations of the mass spectra of ion series [7]; this is determined by the nature of the chosen hypothesis (solution of problems of group mass-spectrometric identification of organic compounds).

Comparison of temperature regimes for the gas-chromatographic analysis of various organic compounds. The next most complex example of large-scatter data arrays can be represented as either a one-dimensional or a two-dimensional array. These are the gas-chromatographic separation parameters of various analytes with temperature programming, namely, the initial temperature (T_0 , °C), the duration of the initial isothermal step ($t(T_0)$, min), and the heating rate (r , K/min). These data from various sources of information are available on the website of the US

National Institute of Standards and Technology (NIST) (<http://webbook.nist.gov>) [11], where they are given for all compounds characterized by retention indices. Any compounds can be chosen to illustrate the features of such information; however, to reduce the volume of discussion, we restrict ourselves to five analytes: nitrobenzene, benzonitrile, 2-chlorophenol, 1,4-dimethoxybenzene, and 1,3,5-trichlorobenzene. Table 2 summarizes some of the values of T_0 , $t(T_0)$, and r for them.

As in the previous example, noticeable variations in all characteristics of the programming modes attract attention. In this case, the standard deviations of the parameters T_0 and r do not exceed the corresponding average values, which cannot be said about the duration of the initial isothermal section (r). From the general point of view, one could assume that the initial temperatures T_0 could be somehow related to the chemical nature of analytes, for example, to the values of their retention indices. However, a check of the RI– T_0 relationship shows that these parameters weakly correlate with each other (Fig. 1) and we can only consider a weakly pronounced tendency to increase the initial temperatures of programming regimes for analytes with higher retention indices. This is especially noticeable if the values of T_0 , $t(T_0)$, and r are not preliminarily averaged for each compound (Fig. 1a), but points corresponding to all values of these parameters are plotted on a graph (Fig. 1b). The correlation coefficient (R) in case (b) is only 0.075, which confirms the absence of a linear relationship between the variables.

If so, the overall average values and standard deviations of the analytical parameters given in the last row of Table 2 can be calculated and used instead of the average values of T_0 , $t(T_0)$, and r for each of the analytes. Because approximately 95% of the sample values should fall within $\pm 2s_{RI}$ intervals, these data correspond to a range of T_0 from 10 to 100°C (no outliers), $t(T_0)$ to 8 min (one outlier, 10 min), and r to 11 K/min (no outliers). The coefficients of asymmetry (A) of all these parameters ranged from 1.6 to 3.5 (the last row of Table 2) for the same reason as in the previous case. Hence, an unexpected conclusion follows: when choosing temperature programming conditions for various compounds, the values of T_0 , $t(T_0)$, and r can be specified in wide ranges, which do not affect the separation of analytes. The most that can be affected by a suboptimal choice of separation conditions (for example, too small values of T_0 , large $t(T_0)$, and low r) is an unjustified increase in the duration of the analysis. Sometimes, the expression *development of a procedure* (for chromatographic analysis or, more precisely, separation) can be used as applied to the choice of separation conditions. However, taking into account the conclusion on the relatively large variations in the parameters T_0 , $t(T_0)$, and r , it should be recognized that it would be more correct to speak about the choice of certain values from the allowable ranges.

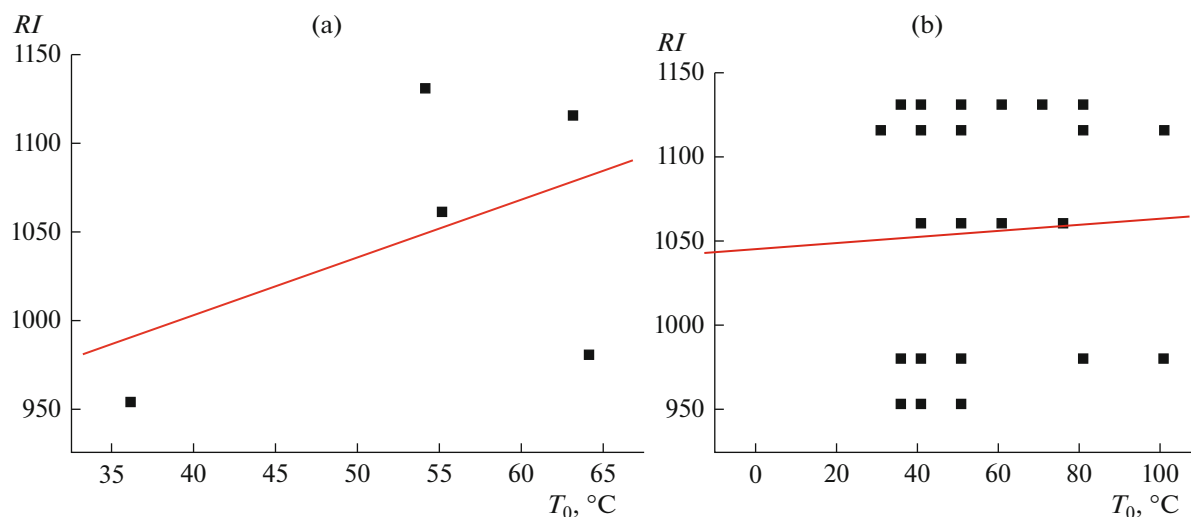


Fig. 1. Graph illustrating the absence of a linear dependence of the retention indices (RIs) of analytes on the initial temperatures of the temperature-programming mode (T_0) chosen for their gas-chromatographic separation: (a) based on the average values of T_0 and (b) according to all data in Table 2. The parameters of the linear regression $RI = aT_0 + b$ are $a = 0.18 \pm 0.45$, $b = 1047 \pm 28$, $R = 0.053$, and $S_0 = 73$.

Characterization of the toxicity (LD_{50}) of homologues based on the example of C_3 – C_{12} 1-alkanols.

The following example can also be considered as both one-dimensional and two-dimensional data sets. For the LD_{50} values of specific compounds, conventional statistical processing is formally possible (if this is consistent with the accepted hypothesis), and a comparison of such data for several homologues turns the array into a two-dimensional array when the least squares method is necessary to calculate the parameters of regression equations. The toxicity of various chemical compounds is one of their most important characteristics, but its determination (especially in vivo) is a very long, laborious, and expensive operation. Numerous attempts are known to theoretically estimate the toxicity parameters (LC_{50} or LD_{50}); however, we restrict ourselves to a few references [66–70] because this is not the subject of this work. At the same time, it is important to note that the calculation methods, one way or another, are based on experimental data, but if their scatter is large, the low accuracy of all estimates inevitably follows from this [63].

Table 3 shows data on the toxicity (LD_{50} , mg/kg, oral) of C_3 – C_{12} 1-alkanols for various warm-blooded animals (rat, rabbit, mouse, hamster, dog, (without breed specification), and bird (the same) and one value for humans (1-propanol)). Here, the significant scatter of the data attracts attention. If we try to average the values of LD_{50} for C_3 – C_6 alcohols (which were characterized in most detail) separately, we obtain estimates for which the coefficients of variation are 45–80% (shown in Table 3). In addition, it is noteworthy that the data samples for C_4 and C_5 alcohols are almost symmetrical; the asymmetry factors for them

are close to unity: 0.90 and 0.97, respectively. The situation with data scatter did not fundamentally change when we confined ourselves to only one species of animals (rats were represented in most detail). For this reason, we can subsequently consider the entire set of objects without any additional subdivision. Table 3 uses the same principle of citing literary sources as that reported previously [71]: references are given to specific works, but sometimes the SDS pages of companies producing particular chemical reagents available on the Internet are mentioned.

This example illustrates the paradoxes that can result from the interpretation of data characterized by high variability. If we calculate the arithmetic mean values of LD_{50} for each of the C_3 – C_{12} alcohols, they correspond to a smooth sinusoidal curve with two extremums in the graph (Fig. 2a). If we consider the entire set of data without preliminary processing (Fig. 2b), their graphical representation illustrates a significant scatter. The authors cannot fail to note the fact that the opinions of specialists regarding the possibility and impossibility of characterizing the toxicity of homologues by average values are divided approximately in equal proportions. However, it is important to emphasize that this choice is also a consequence of one or another initial hypothesis. Considering that the initial values of LD_{50} for each of the alcohols were obtained for different animals, their averaging is unacceptable. If we neglect such information, it is acceptable to average the data for not only each of the homologues but also for their entire set. In this case, the entire set of LD_{50} values can be characterized by the overall mean value and its standard deviation: (2600 ± 1600) mg/kg; on this basis, we can conclude that the available data do not confirm the dependence of the

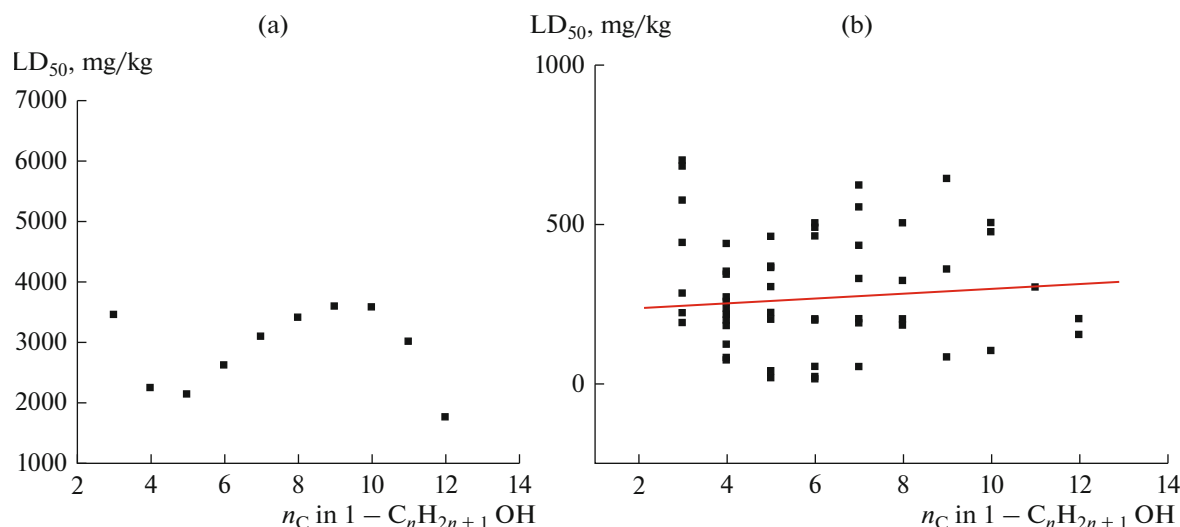


Fig. 2. (a) Graphical illustration of the dependence of the arithmetic mean LD_{50} values of each alcohol on the number of carbon atoms in 1-alkanols; (b) the same for all initial data given in Table 3. The parameters of the linear regression (trend line) $LD_{50} = an_C + b$ are $a = 78 \pm 73$, $b = 2200 \pm 400$, $R = 0.10$, and $S_0 = 1600$.

toxicity of C_3 – C_{12} 1-alkanols on the number of carbon atoms in the molecules. This conclusion is of a general nature: if the standard deviation of the coefficient a for the linear regression $y = ax + b$ exceeds its absolute value, $s_a > |a|$, usual data averaging should be preferred.

Selection of sample preparation operations in the determination of drugs in blood plasma depending on their hydrophobic–hydrophilic properties. The last considered example of processing and interpreting data characterized by a large scatter is the most complex. First, it refers to two-dimensional arrays, and one of the variables should be chosen artificially (by introducing the ranks of sample preparation operations). The second factor is the relevance of the problem under consideration. The determination of drugs in biological fluids is a very labor-intensive task primarily due to the complexity of a matrix and, as a result, a sample preparation stage. Because reversed phase HPLC is the preferred analytical method for solving such problems, it is necessary to ensure not only the preconcentration of target analytes at this stage but also the removal of interfering components, primarily, proteins. Sample preparation most often includes the following operations: liquid–liquid and solid-phase extraction, precipitation of protein components, ultrafiltration, ion exchange, and, less often, other operations. The selection and optimization of sample preparation operations (or their combinations) is still carried out on the basis of general concepts of the nature of the compounds to be determined. This results in a large amount of time, which could be minimized if it were possible to relate the nature of sample preparation operations to the physicochemical characteristics of the analytes.

Of these characteristics, the hydrophobicity factor $\log P$ deserves attention in the first place. However, the experimental values of $\log P$ are known not for all characterized compounds. They can be replaced by calculated estimates, in our case, calculated using the ChemAxon software [19]. When such estimates are obtained, the values of $\log P(\text{ChemAxon}) < -2$ can be excluded from consideration because there are no examples of the use of LLE for such hydrophilic compounds. Figure 3 illustrates checking the equivalence of calculated (indicated in Table 4 for some drugs) and experimental values (13 pairs of values) and shows that the coefficient a of the linear regression $\log P(\text{expt}) = a \log P(\text{ChemAxon}) + b$ is only slightly lower than unity (0.89 ± 0.09), while the coefficient b is statistically insignificant (0.27 ± 0.31). Therefore, the experimental and calculated values of $\log P$ can be considered equivalent in some approximation. However, the average accuracy of the resulting estimates is not high because the value of S_0 for the linear regression shown in Fig. 3 is 0.69. Moreover, with the use of the calculated values of $\log P$, it is undesirable to use the values of $\log P$ calculated in other ways (for example, using the ACD software).

Table 4 compares the most important sample preparation conditions for 51 drug preparations; molecular weight, CAS no., $\log P(\text{ChemAxon})$, and a fraction of drug bound to plasma proteins are specified for each particular drug. A reference to the original publication is provided for each version of sample preparation; the absence of a reference (10 preparations) indicates that the procedure was developed and validated at JSC Biokad.

First of all, let us comment on such a characteristic as the fraction of analytes bound to plasma proteins

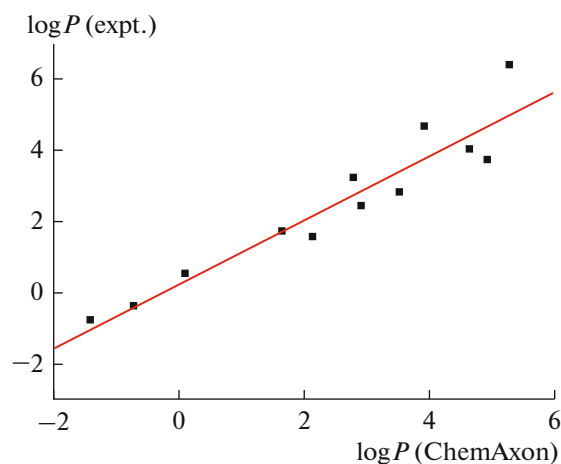


Fig. 3. Correlation between the experimental and calculated (ChemAxon) values of $\log P$. The parameters of the linear regression $\log P(\text{expt}) = a \log P(\text{ChemAxon}) + b$ are $a = 0.89 \pm 0.09$, $b = 0.28 \pm 0.31$, $R = 0.944$, and $S_0 = 0.69$.

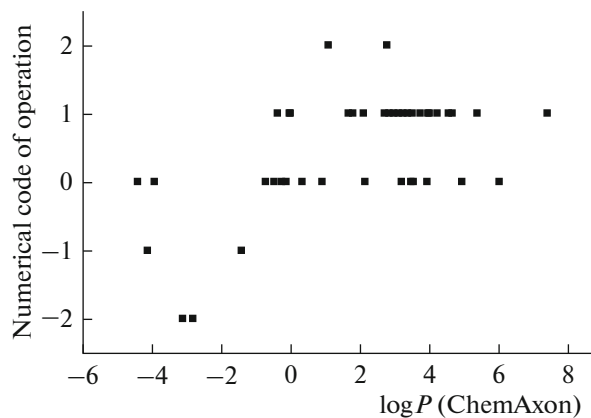


Fig. 4. Graphical illustration of a correlation of the numerical codes of the main sample preparation operations in the determination of drug traces in blood plasma with the values of $\log P$ (ChemAxon).

(varies from 2 to 99%), which directly affects their detection limits. Checking a possible relationship of this characteristic with a sample preparation operation such as protein precipitation shows that there is no significant correlation here. The average value of the bound fraction of the analyte with the use of precipitation is $(89 \pm 16)\%$, and it is $(70 \pm 33)\%$ in the absence of this operation (according to the data of Table 4). Thus, the case in point is only a weakly expressed trend, which can be ignored.

For the subsequent interpretation of data, it is necessary to assign tentative codes (ranks) to various sample preparation operations, which will allow us to apply a method somewhat similar to Spearman's rank

Sample preparation operation	SPE	LLE
Numerical code	+2	+1

As in the previous examples, the results of checking the possible relationship between the values of $\log P$ and the proposed numerical codes can be most clearly presented graphically (Fig. 4). From Fig. 4, it follows that, firstly, two versions of extraction (LLE and SPE) are much more popular methods of sample preparation compared to the others; this is mainly due to the hydrophobic properties of the characterized drugs. However, even with such an uneven distribution of points as that in Fig. 4, we can conclude that there are no examples of the use of ultrafiltration and ion exchange at $\log P > -1$ or (alternative wording) only LLE or SPE are used at the stage of sample preparation. The LLE and SPE operations are fundamentally inseparable, and a choice between them is determined primarily by the availability of appropriate equipment,

correlation. The tentative zero value of a numerical code can be assigned to the protein precipitation operation because it is least of all related to the hydrophobic–hydrophilic properties of the target analytes by its very nature. SPE (code, +2) is used for the preconcentration of the most hydrophobic compounds, and the code can be equated to +1 in the case of less hydrophobic compounds, when the use of LLE is acceptable (the solvents can be varied). Such an operation as ion exchange is applicable only to compounds that exist in the ionic form (the most hydrophilic compounds; code, -2). Then, the ultrafiltration procedure receives the code (-1). Thus, we obtain the following set of variables for rank correlation:

Precipitation	Ultrafiltration	Ion exchange
0	-1	-2

materials, and solvents. Note that LLE at $\log P < 0$ is fundamentally possible but much more difficult to perform, and there are no examples of the use of this extraction in the region of $\log P < -1$ among the drugs under consideration.

CONCLUSION

Thus, the examples of data arrays with a large levels of variations are widespread. The most difficult are the examples of variability due to differences in the nature of the objects themselves. In these cases, data interpretation is complicated by the influence of the accepted analytical hypotheses. Such data arrays with high degrees of uncertainty are often excluded from consideration, and this is not always justified. The processing

of these data has its own characteristics. As a rule, one-dimensional sets $\{x_i\}$ are characterized by high asymmetry. Large values of relative standard deviations $\delta_i = s_x/\langle x \rangle$ can be interpreted as the absence of the need for precise control of the values of variables x , which can be replaced by a choice from a wide range of their possible values. In contrast to the rules of conventional statistical processing, the condition $s_x > \langle x \rangle$ does not mean that the average value $\langle x \rangle$ is zero but that some of the values $\{x_i\}$ of the sample can be equal to zero. If the standard deviation of the coefficient a for the linear regression $y = ax + b$ exceeds its value ($s_a > |a|$) as a result of processing some data arrays by the least squares method, the usual averaging of the data with an estimate of $\langle x \rangle \pm s_x$ should be preferred. The probabilistic nature of the conclusions based on data with high degrees of uncertainty, as a rule, is higher than that for data with normal reproducibility. At the same time, important conclusions can be drawn even from the very facts of high variability of variables. For example, Zenkevich and Deruish [72], who characterized the dependence of the retention indices of analytes in reversed phase HPLC on the methanol content of the eluant (dRI/dc), found that these parameters, in contrast to the retention indices themselves, do not correlate with the values of $\log P$; this fact makes it possible to exclude such a correlation from the subsequent consideration.

CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest.

OPEN ACCESS

This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

REFERENCES

1. *Analiticheskaya khimiya* (Analytical Chemistry), vol. 3: *Khimicheskii analiz* (Chemical Analysis), Moskvina, L.N., Ed., Moscow: Akademiya, 2010.
2. *Guideline on Bioanalytical Method Validation*, London: Eur. Med. Agency, 2011.
3. The US Pharmacopeia 35, Section Chromatographia/Physical Tests/System Suitability, New York: The USP Convention, 2012, p. 262.
4. On Approval of the Rules for Conducting Bioequivalence Studies of Medicinal Products within the Framework of the Eurasian Economic Union, no. 85, Astana: Evraz. Ekon. Komissiya, 2016.
5. Adams, F. and Thomas, J.M., *Philos. Trans. R. Soc., A*, 1981, vol. 305, p. 509. <https://doi.org/10.1098/rsta.1982.0048>
6. Becker, S. and Dietze, H.-J., *Int. J. Mass Spectrom. Ion Processes*, 2000, vol. 197, nos. 1–3, p. 1. [https://doi.org/10.1016/S1387-3806\(99\)00246-8](https://doi.org/10.1016/S1387-3806(99)00246-8)
7. Zenkevich, I.G. and Ioffe, B.V., *Interpretatsiya mass-spektrov organicheskikh soedinenii* (Interpretation of Mass Spectra of Organic Compounds), Leningrad: Khimiya, 1986.
8. Lochmüller, C.H. and Hui, M., *J. Chromatogr. Sci.*, 1998, vol. 36, p. 11.
9. Hanai, T., Koizumi, K., and Kinoshita, T., *J. Liq. Chromatogr. Relat. Technol.*, 2000, vol. 23, p. 363. <https://doi.org/10.1081/JLC-100101457>
10. Yashtold-Govorko, V.A., *Fotos'emka i obrabotka* (Photography and Processing), Moscow: Iskusstvo, 1964.
11. NIST 20 (2020) Mass Spectral Library (NIST/EPA/NIH EI MS Library, 2020 Release). Software/Data Version; NIST Standard Reference Database no. 69, Gaithersburg: Natl. Inst. Standards Technol., 2020. <http://webbook.nist.gov>. Accessed December 2021.
12. Patochka, J. and Kuca, K., *Mil. Med. Sci. Lett.*, 2012, vol. 81, no. 4, p. 142. <https://doi.org/10.31482/mml.2012.022>
13. *n*-Butyl Alcohol, CAS no. 71-36-3, SIDS Initial Assessment Report, Bern: UNEP, 2001.
14. Butanols: Four Isomers, Environment Health Criteria no. 65, New York: World Health Org., 1987.
15. *GOST* (State Standard) 30333-2007: *Chemical Production Safety Passport. General Requirements*, Moscow: Izd. Standartov, 2008.
16. Substance Evaluation Conclusion for Butan-1-ol. EC no. 200-751-6, Budapest: Ministry for Human Capacities, 2018.
17. Case study on the use of integrated approaches for testing and assessment of 90-day rat oral repeated-dose toxicity for selected *n*-alkanols: read across, OECD Environment. Health and Safety Publ. no. 273, Paris: Organization for Economic Co-operation and Development, 2017.
18. Wypych, A. and Wypych, G., *Databook of Solvents*, Amsterdam: Elsevier, 2019, 2nd ed.
19. Yanagimachi, N., Obara, N., Sakata-Yanagimoto, M., Chiba, S., Doki, K., and Homma, M., *Biomed. Chromatogr.*, 2021, vol. 35, no. 5, e5049. <https://doi.org/10.1002/bmc.5049>
20. Bahrami, G., Mohammadi, B., Mirzaeei, S., and Kiani, A., *J. Chromatogr. B: Anal. Technol. Biomed. Life Sci.*, 2005, vol. 826, nos. 1–2, p. 41. <https://doi.org/10.1016/j.jchromb.2005.08.008>

21. Said, R., Arafat, B., and Arafat, T., *J. Chromatogr. B: Anal. Technol. Biomed. Life Sci.*, 2020, vol. 1149, p. 122. <https://doi.org/10.1016/j.jchromb.2020.122154>
22. Qian, J., Wang, Y., Chang, J., Zhang, J., Wang, J., and Hu, X., *J. Chromatogr. B: Anal. Technol. Biomed. Life Sci.*, 2011, vol. 879, no. 9, p. 662. <https://doi.org/10.1016/j.jchromb.2011.01.039>
23. Vanwelkenhuysen, I., de Vries, R., Timmerman, P., and Verhaeghe, T., *J. Chromatogr. B: Anal. Technol. Biomed. Life Sci.*, 2014, vol. 958, p. 43. <https://doi.org/10.1016/j.jchromb.2014.02.028>
24. Patel, B.N., Sharma, N., Sanyal, M., and Shrivastav, P.S., *J. Chromatogr. Sci.*, 2010, vol. 48, no. 1, p. 35. <https://doi.org/10.1093/chromsci.48.1.35>
25. Casas, M., Hansen, M., Krogh, K., Styrihave, B., and Björklund, E., *J. Chromatogr. B: Anal. Technol. Biomed. Life Sci.*, 2014, vol. 962, p. 109. <https://doi.org/10.1016/j.jchromb.2014.02>
26. Hammad, M., Kamal, A., Kannouma, R., and Mansour, F., *J. Chromatogr. Sci.*, 2021, vol. 59, no. 3, p. 297. <https://doi.org/10.1093/chromsci/bmaa096>
27. Qiu, F., Gu, Y., Wang, T., Gao, Y., Li, X., Gao, X., and Cheng, S., *Biomed. Chromatogr.*, 2016, vol. 30, p. 962. <https://doi.org/10.3390/pharmaceutics10040221>
28. Wong, A., Xiang, X., Ong, P., Mitchell, E., Syn, N., Wee, I., Kumar, A., Yang, W., Sethi, G., Goh, B., Ho, P., and Wang, L., *Pharmaceutics*, 2018, vol. 10, no. 4, p. 221. <https://doi.org/10.3390/pharmaceutics10040221>
29. Bahrami, G. and Mohammadi, B., *Clin. Chim. Acta*, 2006, vol. 370, nos. 1–2, p. 185. <https://doi.org/10.1016/j.cca.2006.02.017>
30. Mabrouk, M., Soliman, S., El-Agizy, H., and Mansour, F., *J. Chromatogr. B: Anal. Technol. Biomed. Life Sci.*, 2020, vol. 1136. <https://doi.org/10.1016/j.jchromb.2019.121932>
31. Aravagiri, M., Ames, D., Wirshing, W.C., and Marder, S.R., *Ther. Drug Monit.*, 1997, vol. 19, no. 3, p. 307. <https://doi.org/10.1097/00007691-199706000-00011>
32. Yuan, L., Jiang, H., Ouyang, Z., Xia, Y., Zeng, J., Peng, Q., Lange, R., Deng, Y., Arnold, M., and Aubry, A., *J. Chromatogr. B: Anal. Technol. Biomed. Life Sci.*, 2013, vol. 921, p. 81. <https://doi.org/10.1016/j.jchromb.2013.01.029>
33. Lobo, K., Gross, A., Ray, J., and McLachlan, A., *J. Chromatogr. B: Anal. Technol. Biomed. Life Sci.*, 2005, vol. 823, no. 2, p. 115. <https://doi.org/10.1016/j.jchromb.2005.06.009>
34. Peigné, S., Chhun, S., Tod, M., Rey, E., Rodrigues, C., Chiron, C., Pons, G., and Jullien, V., *Clin. Pharmacokinet.*, 2018, vol. 57, no. 6, p. 739. <https://doi.org/10.1007/s40262-017-0592-7>
35. Takahashi, R., Imai, K., and Yamamoto, Y., *Jpn. J. Pharm. Health Care Sci.*, 2015, vol. 41, p. 643. <https://doi.org/10.5649/jjphcs.41.643>
36. Rouini, M., Ardakani, Y., Moghaddam, K., and Solatani, F., *Talanta*, 2008, vol. 75, no. 3, p. 671. <https://doi.org/10.1016/j.talanta.2007.11.060>
37. Minkin, P., Zhao, M., Chen, Z., Ouwerkerk, J., Gelderblom, H., and Baker, S.D., *J. Chromatogr. B: Anal. Technol. Biomed. Life Sci.*, 2008, vol. 874, p. 84. <https://doi.org/10.1016/j.jchromb.2008.09.007>
38. Wang, L., Goh, B., Grigg, M., Lee, S., Khoo, Y., and Lee, H., *Rapid Commun. Mass Spectrom.*, 2003, vol. 17, no. 14, p. 1548. <https://doi.org/10.1002/rcm.1091>
39. Bahrami, G. and Mohammadi, B., *J. Chromatogr. B: Anal. Technol. Biomed. Life Sci.*, 2007, vol. 857, no. 2, p. 322. <https://doi.org/10.1016/j.jchromb.2007.07.044>
40. Yasu, T., Sugi, T., Momo, K., Hagihara, M., and Yasui, H., *Biomed. Chromatogr.*, 2021, vol. 35, no. 4, e5028. <https://doi.org/10.1002/bmc.5028>
41. Perez, H., Boram, S., and Evans, C., *Anal. Methods*, 2015, vol. 7, no. 2, p. 723. <https://doi.org/10.1039/C4AY02599G>
42. Mohamed, F., Ali, M., Marwa, F.B., Rageh, A., and Mostafa, A., *Microchem. J.*, 2019, vol. 146, p. 414. <https://doi.org/10.1016/j.microc.2019.01.031>
43. Wang, L., Wang, J., Zhang, J., Jiang, Q., Zhao, L., and Zhang, T., *Arab. J. Chem.*, 2019, vol. 12, no. 8, p. 4775. <https://doi.org/10.1016/j.arabjc.2016.09.016>
44. Charbe, N., Baldelli, S., Cozzi, V., Castoldi, S., Cattaneo, D., and Clementi, E., *J. Pharm. Anal.*, 2016, vol. 6, no. 6, p. 396. <https://doi.org/10.1016/J.JPHA.2016.05.008>
45. Dharmalingam, S.R., Ramamurthy, S., Chidambaram, K., and Nadaraju, S., *Trop. J. Pharm. Res.*, 2014, vol. 13, no. 3, p. 409. <https://doi.org/10.4314/tjpr.v13i3.15>
46. Lourenco, D., Sarraguca, M., Alves, G., Coutinho, P., Araujo, A., and Rodrigues, M., *Anal. Methods*, 2017, vol. 9, no. 40, p. 5910. <https://doi.org/10.1039/C7AY01912B>
47. Majnooni, M.B., Mohammadi, B., Jalili, R., and Bahrami, G.H., *Indian J. Pharm. Sci.*, 2012, vol. 74, no. 4, p. 360. <https://doi.org/10.4103/0250-474X.107073>
48. Kaushik, K., Sripuram, V., Bedada, S., Reddy, N., Priyadarshini, G., and Devarakonda, K., *Clin. Res. Reg. Affairs*, 2010, vol. 27, no. 1, p. 1. <https://doi.org/10.3109/10601330903490462>
49. Naik, K. and Nandibewoor, S., *J. Anal. Chem.*, 2013, vol. 68, no. 12, p. 1212. <https://doi.org/10.7868/S0044450213120049>
50. Madej, K., Paprotny, L., Wianowska, D., Kasprzyk, J., Herman, M., and Piekoszewski, W., *Biomed. Chromatogr.*, 2021, vol. 35, no. 3, e5002. <https://doi.org/10.1002/bmc.5002>
51. Hemasree, S., Sumadhuri, B., and Murthy, T., *J. Adv. Pharm. Educ. Res.*, 2013, vol. 3, no. 3, p. 187.
52. Galmier, M., Frasey, A., Bastide, M., Beyssac, E., Petit, J., Aiache, J., and Lartigue-Mattei, C., *J. Chromatogr. B: Biomed. Sci. Appl.*, 1998, vol. 720, nos. 1–2, p. 239. [https://doi.org/10.1016/S0378-4347\(98\)00443-5](https://doi.org/10.1016/S0378-4347(98)00443-5)
53. Dasht Bozorg, B., Goodarzi, A., Fahimi, F., Tabarsi, P., Shahsavari, N., Kobarfard, F., and Dastan, F., *Iran. J. Pharm. Res.*, 2019, vol. 18, no. 4, p. 1735. <https://doi.org/10.22037/ijpr.2019.1100849>
54. Hilhorst, M., Hendriks, G., van Hout, M., Sillen, H., and van de Merbel, N., *Bioanalysis*, 2011, vol. 3, no. 14,

- p. 1603.
<https://doi.org/10.4155/bio.11.140>
55. Anders, N., Wanjiku, T., He, P., Azad, N., and Rudek, M., *Biomed. Chromatogr.*, 2016, vol. 30, no. 3, p. 494.
<https://doi.org/10.1002/bmc.3562>
56. Assessment report for vidaza (international nonproprietary name: Azacytidine) no. 593162, Eur. Med. Agency, 2008.
57. D'Avolio, A., Peila, E., Simiele, M., Pensi, D., Baietto, L., Cusato, J., Cinnirella, G., de Rosa, F., and di Perri, G., *Ther. Drug Monit.*, 2013, vol. 35, no. 6, p. 853.
<https://doi.org/10.1097/FTD.0b013e31829403b1>
58. Bergman, J., Harvill, L., Hawkins, S., Sladky, K., and Cox, S., *Biomed. Chromatogr.*, 2021, vol. 35, no. 7, e5104.
<https://doi.org/10.1002/bmc.5104>
59. Dincel, D., Sagirli, O., and Topcu, G., *J. Chromatogr. Sci.*, 2020, vol. 58, no. 2, p. 144.
<https://doi.org/10.1093/chromsci.bmz087>
60. Paal, M., Zoller, M., Schuster, C., Vogeser, M., and Schutze, G., *J. Pharm. Biomed. Anal.*, 2018, vol. 152, p. 102.
<https://doi.org/10.1016/j.jpba.2018.01.031>
61. Roth, T., Fiedler, S., Mihai, S., and Parsch, H., *Biomed. Chromatogr.*, 2017, vol. 31, no. 5, e3880.
<https://doi.org/10.1002/bmc.3880>
62. RUGBANK Online. <https://go.drugbank.com/drugs>. Accessed November 2021.
63. Zenkevich, I.G., *Russ. J. Phys. Chem. A*, 2021, vol. 95, p. 894.
<https://doi.org/10.1134/S0036024421040294>
64. Linnik, Yu.V., *Metod naimen'shikh kvadratov i osnovy teorii obrabotki nablyudenii* (The Method of Least Squares and the Fundamentals of the Theory of Processing Observations), Moscow: Fizmatgiz, 1958.
65. Zenkevich, I.G., *Vestn. St. Petersburg. Gos. Univ., Ser. Fiz.-Khim.*, 1998, no. 2, p. 84.
66. Wu, K. and Wei, G.-W., *J. Chem. Soc. Int. Model.*, 2018, vol. 58, no. 2, p. 520.
<https://doi.org/10.1021/acs.jvum.7b00558>
67. Pu, L., Naderi, M., Liu, T., Wu, H.-C., Mukhopadhyay, S., and Brylinski, M., *BMC Pharmacol. Toxicol.*, 2019, vol. 20, 2.
<https://doi.org/10.1186/s40360-018-0282-6>
68. Semenova, E. and Lazic, S.E., *Comput. Toxicol.*, 2020, vol. 16, p. 100.
<https://doi.org/10.1016/j.comtox.2020.100133>
69. Xia, X., *Cell Biol. Toxicol.*, 2020, vol. 36, p. 591.
<https://doi.org/10.1007/s10565-020-09552-2>
70. Rim, K.-T., *Toxicol. Environ. Health Sci.*, 2020, vol. 15, p. 1.
<https://doi.org/10.1007/s13530-020-00056-4>
71. Zenkevich, I.G., Nikitina, D.A., and Deruish, A., *J. Anal. Chem.*, 2021, vol. 76, p. 493.
<https://doi.org/10.1134/S1061934821040146>
72. Zenkevich, I.G. and Deruish, A., *Anal. Kontrol'*, 2022, vol. 26, no. 1, p. 41.
<https://doi.org/10.15826/analitika.2022.26.1.004>

Translated by V. Makhlyarchuk