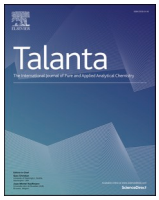
Advanced methods for assessment of risks of false decisions in analytical chemistry (testing) laboratories – A review

Talanta 294 (2025) 128208

Contents lists available at ScienceDirect

Talanta

journal homepage: www.elsevier.com/locate/talanta



Advanced methods for assessment of risks of false decisions in analytical chemistry (testing) laboratories – A review

Ilya Kuselman [a,*], Francesca R. Pennecchi [b], D. Brynn Hibbert [c], Angelique Botha [d], Tamar Gadrich [e], Anastasia A. Semenova [f]

[a] Independent Consultant on Metrology, 4/6 Yarehim St., 7176419, Modiin, Israel
[b] Istituto Nazionale di Ricerca Metrologica (INRIM), Strada delle Cacce 91, 10135, Turin, Italy
[c] School of Chemistry, UNSW Sydney, Sydney, NSW, 2052, Australia
[d] National Metrology Institute of South Africa (NMISA), Private Bag X34, Lynnwood Ridge, 0040, Pretoria, South Africa
[e] Braude College of Engineering, Department of Industrial Engineering and Management, P.O. Box 78, 51 Snunit St., 2161002, Karmiel, Israel
[f] V.M. Gorbatov Federal Research Center for Food Systems, 26 Talalikhina St., 109316, Moscow, Russia

ARTICLE INFO

Keywords:
Risk assessment
False decisions
Analytical chemistry
Testing laboratory
Human errors
Measurement uncertainty
Conformity assessment

ABSTRACT

There are two groups of decision-making risks in an analytical chemistry (testing) laboratory directly influencing quality of measurement/test results. One group consists of the risks of false decisions caused by human errors in performing a test. The second group of risks is from the erroneous interpretation of test results, due to measurement uncertainty, judged against the specification/tolerance limits (conformity assessment). Basic concepts of advanced methods for the assessment of risks of false decisions in an analytical chemistry laboratory that have been developed in the last decade are reviewed in the present paper.

1. Introduction

Risk management has a long history, recorded, perhaps for the first time, in the Tigris-Euphrates valley about 3200 BCE by the Asipu group of consultants for risky, uncertain, or difficult decisions [1]. The importance of risk management is also recognized in the Torah and Bible. For example, it is written in the Bible: "A prudent person foresees danger and takes precautions. The simpleton goes blindly on and suffers the consequences" [2]. Modern risk management started in the 1950s, after the Second World War [3]. About the same time the words "risk", "danger", and "hazard" (used earlier as near neighbours) moved apart because of a semantic drift of "risk" toward forecasting and prevention [4]. Risk management, as a scientific field, is about 40 years old [5].

Nowadays, there are three main documents for risk management published by the International Organization for Standardization (ISO) and Electrotechnical Commission (IEC): vocabulary [6], guidelines [7], and risk assessment techniques [8]. Risk is defined here as an effect of uncertainty on results to be achieved, expressed in terms of risk sources, potential events, their likelihood (probability) and consequences. The documents [6–8] provide a common approach for the management of any type of risk (including its minimization) based on risk assessment and are not industry- or sector-specific. They can be customized to any organization and applied to any activity, including decision-making. However, the authors of the standards [9] emphasize that in practice, the concepts and terminology need to be adapted to the field or discipline of application, to avoid misinterpretation, misrepresentation, or misuse. Therefore, several guidelines for specific fields have been developed, for example, the quality risk management guidelines for industry by the International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use (ICH) [10], and the ISO standard for application of risk management to medical laboratories [11].

Risk assessment is the three-step process of 1) risk identification – finding, recognizing, and describing; 2) analysis – understanding the nature of the risk and the determination of its level/magnitude; and 3) evaluation – comparing the results of risk analysis with risk criteria to determine whether the risk is acceptable or tolerable [6–8]. Since risk assessment is the basic part of risk management, every guide on the topic includes a description of risk assessment, e.g., in the guidance for a disaster risk management [12]. Law and regulations are applied for risk

* Corresponding author. 4/6 Yarehim St., Modiin, 7176419, Israel.
E-mail address: ilya.kuselman@bezeqint.net (I. Kuselman).

https://doi.org/10.1016/j.talanta.2025.128208
Received 15 February 2025; Received in revised form 18 April 2025; Accepted 22 April 2025
Available online 24 April 2025

assessment of agrochemicals and in similar fields important for public health [13].

There are also specific guidelines on risk assessment, for example, the guidance on risk assessment at work [14]. Risk assessment at laboratory work within the scope of occupational health and safety of workers, financial sustainability, and protection of the environment, is required [15,16]. This task, especially important for academic and university chemical laboratories teaching students [17–19], is still challenging [20–22].

Risk assessment of errors in clinical and medical laboratories aiming for correct medical diagnosis and patient safety is developed intensively, considering pre- and post-analytical steps (phases) of testing, and their automatization/robotization [23–25]. For example, the majority of human errors reported in a study of clinical and pathology laboratories in a selection of Iranian hospitals [26] happened at the patient's admission stage, including incorrect recording of patient information and registration of the requested tests. The reason was that the great number of patients referred to the hospital laboratories exceeded the laboratory's normal throughput. An irregular (individual) diagnostic sample can compromise any analytical measurement procedure [27]. The need for such study is widely recognized [28,29] and considered in corresponding standards [11,30].

In an analytical chemistry (testing) laboratory, assessment of decision-making risks that directly influence the quality of measurement/test results is an important part of the laboratory's competence required for its accreditation [31]. There are risks of false decisions in 1) performing a chemical analysis/test, and 2) interpretation of the measurement/test results when comparing them to the specification/tolerance limits of a test item for conformity assessment. The risks in performing a chemical analysis are caused by human errors, while the risks of erroneous conformity assessment arise due to measurement uncertainty.

The joint guide on the risks from human errors in a laboratory performing chemical analysis was developed by International Union of Pure and Applied Chemistry (IUPAC) and Cooperation on International Traceability in Analytical Chemistry (CITAC) [32] using the experience from investigations of accidents in aviation, engineering, security, and other fields. In its turn, this guide was implemented not only for analysis of water, food, geological samples, and environmental pollutants [33], but also for quality control in a medical laboratory [34] and in such unexpected field as military equipment testing [35].

The topic of risks of false decisions in conformity assessment of a substance, material, or object due to measurement uncertainty was initially addressed in the IUPAC/CITAC Guide [36] on investigating out-of-specification test results of a single component content. It was further developed *ibid* in the guide [37] for multicomponent objects, when test results may be correlated, and in the guide [38] for multicomponent objects under a mass balance constraint. The guides [36–38] related to the evaluation of measurement data by the Joint Committee for Guides in Metrology (JCGM), described in the JCGM 106 document [39], were implemented for different kinds of analysis of the chemical composition of a substance, material, or object. Besides, these approaches have been applied recently in several pharmaceutical studies at the Universidade de São Paulo, Brazil [40–42].

Thus, there is a fruitful exchange of information among the different fields of risk assessment, adaptation and use of suitable approaches. In addition to field guidelines, several scattered scientific reports on risk assessment that may be attractive for analytical chemists have been published during the past decade. For example, there are publications on the risks caused by cognitive issues and human errors in forensic investigations [43–45], on principles of risk-decision making in public health [46] and in setting performance requirements for an analytical procedure [47].

A new IUPAC project [48] for harmonization of basic concepts and associated terms applied in advanced methods for the assessment of the risks of false decisions in analytical chemistry laboratories has been recently started. The present paper is a review of the advanced methods that have been developed in the last decade.

2. Risks of false decisions caused by human errors in performing chemical analysis

Despite the development of risk assessment and methods to manage risk, human error is here to stay [49–51] since *errare humanum est* (to err is human). The role of a human being as a part of a measuring system in an analytical chemistry laboratory is discussed in the paper [52] published in memory of the late Paul De Bièvre. It is highlighted there that a measuring system in chemical analysis includes not only measuring instruments and other devices, reagents, and supplies, but also a sampling inspector and/or analyst, making decisions and performing necessary operations. In general, without the human contribution, a measurement cannot be carried out even in a robotic laboratory, which is, anyway, designed and built by specialists in analytical chemistry and other fields.

A human error in chemical analysis is any action, or lack thereof, that leads to exceeding the tolerances of the conditions required for the correct functioning of the measuring/testing (chemical analytical) system. Tolerances of the conditions are, for example, intervals of temperature and pressure values for a sample decomposition, purity of reagents, pH values for an analyte extraction and separation, etc. They are formulated in a standard operating procedure (SOP) of the analysis describing the correct work, based on results of the analytical method validation study.

In a routine analytical laboratory human error may lead to atypical test results of questionable reliability. An important group of atypical results is out-of-specification test results [53] in the pharmaceutical industry, not complying with regulatory, legislation or specification limits in other industries and fields, for example, environmental and food analysis.

Risk of human error is the product of the likelihood of occurrence of such an error and its consequence/severity for the quality of analytical results. Prevention, avoidance or blocking of human errors by a laboratory quality system is not easy as both correct performance and errors follow from the same cognitive processes allowing us to be fast, to respond flexibly to new situations, and to juggle several tasks at once. Both are "two sides of the same theoretical coin." Human errors may occur for various reasons. Many seem trivial for professionals. However, people do make trivial errors in their routine work. Nobody can change human nature. Thus, protection of the quality of an analytical result by managing the risk of human error, to reduce the likelihood of error and mitigate its severity (risk reduction), is an important task for the quality system of any analytical chemistry laboratory. Residual risk of human error, not prevented or blocked by the laboratory quality system, decreases the quality of analytical results and may be interpreted as a source of measurement uncertainty.

Each analytical chemistry laboratory has its facilities, equipment, budget, and staff. However, a laboratory does not have usually a database containing empirical values of likelihood/frequencies of occurrence of human errors and corresponding severities. On the other hand, any expert in a specific chemical analysis has the necessary information accumulated during his/her work. That is why classification, modelling and evaluation of human errors in chemical analysis are discussed in the guide [32] and other publications using expert estimations (judgments).

2.1. Classification

There are nine kinds of human errors, $k = 1, 2, \ldots, K$ ($K = 9$) including seven kinds of commission errors of a sampling inspector and/or an analyst/operator (knowledge-, rule- and skill-based mistakes; and routine, reasoned, reckless, and malicious violations) and two kinds of omission errors (lapses and slips).

These errors may happen at any step of chemical analytical measurement/testing process, $m = 1, 2, \ldots, M$ (location of the error). The

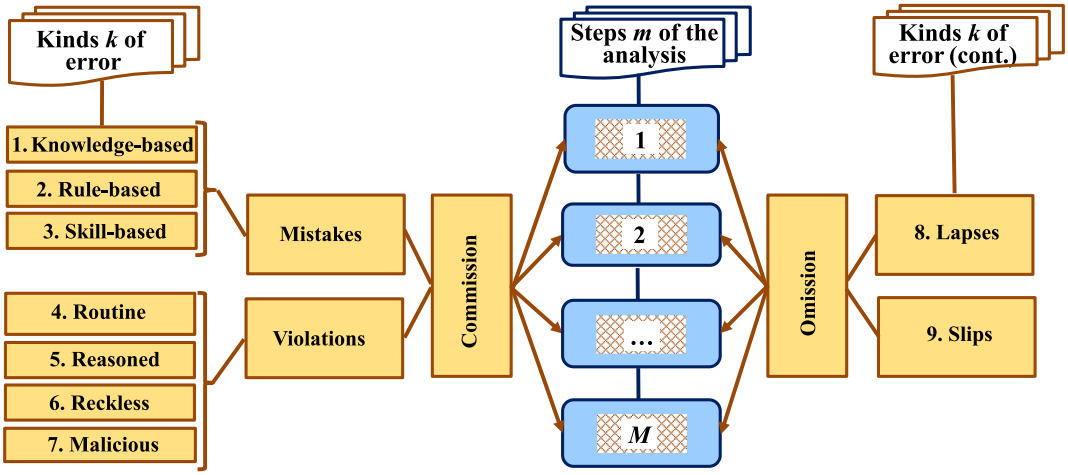


Fig. 1. A map of human errors in chemical analysis. Pointers show links of errors to steps of the analytical process. Nets indicate error scenarios. Modified from Ref. [32].

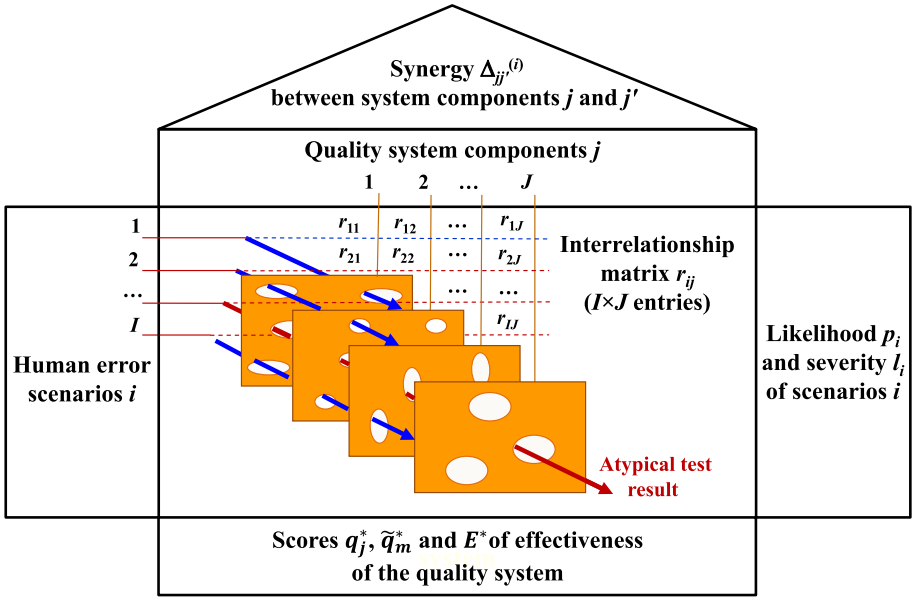


Fig. 2. A laboratory quality system against human errors. Human error scenarios are indicated as rows $i = 1, 2, \ldots, I$. Quality system components/layers are shown as the Swiss cheese slices (columns) $j = 1, 2, \ldots, J$. Bad outcomes blocked by the quality system layers are presented by the blue pointers. Appearance of an atypical test result is depicted by the longest red pointer. Estimates of reduction r_{ij} of likelihood and severity of error scenario i, as the result of interaction between the error and layer j, form the interrelationship matrix here. Modified from Ref. [32]. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

main steps typically are: 1) choice of the chemical analytical method and corresponding SOP, 2) sampling, 3) analysis of a test portion, and 4) calculation of test results and reporting. However, after sampling, sample preparation is required in many chemical analytical methods, including sample freezing, milling and/or decomposition. The chemical analysis may start from extraction of an analyte from a test portion and separation of the analyte from other components of the extract. Analyte identification and confirmation are important. Then only calibration of the measuring system and quantification of the analyte concentration are relevant. On the other hand, choice of an analytical method and SOP may not be necessary in a laboratory where only one method and corresponding SOP are applied or mandated, for a specific task. Many analytical chemistry laboratories are not responsible for sampling, for example crime-scene sampling by forensic police.

The kind of human error and the step of the analysis, in which the error may occur, form the event scenario, $i = 1, 2, \ldots, I$. There are at most $I = K \times M$ scenarios of human errors. Since $K = 9$ here, hence $I = 9M$. These scenarios, put together, generate a map of human errors in chemical analysis shown in Fig. 1.

2.2. Modelling

A Swiss cheese model [54] shown in Fig. 2 is used for characterizing the errors interaction with a laboratory quality system. This model considers the quality system components $j = 1, 2, \ldots, J$ as protective layers against human errors. For example, the system components may be: validation of the measurement/analytical method and formulation of the SOP; training of analysts and proficiency testing; quality control using statistical charts and/or other means; and supervision. Each component has weak points, whereby errors are not prevented, like holes in slices of the cheese. As a rule, the presence of holes in a layer will not lead to system failure, since other layers are able to prevent a

bad outcome. This is shown in Fig. 2 as the blue pointers blocked by the layers. An incident occurs and an atypical test result will appear, when the holes in the layers line up at the same time to permit a trajectory of incident opportunity to pass the system through its defects, as depicted in Fig. 2 by the longest red pointer.

2.3. Quantification

A technique for quantifying human errors in chemical analysis using expert judgments was formulated based on the Swiss cheese model and the house-of-security approach [55]. According to this approach, an expert may estimate the likelihood p_i of scenario i by the following ordinal scale: likelihood of an unfeasible scenario – as $p_i = 0$, weak likelihood – as $p_i = 1$, medium – as $p_i = 3$, and strong (maximal) likelihood – as $p_i = 9$. The expert estimates/judgments on the severity of an error by scenario i, interpreted as the expected loss l_i of quality of the analysis result, are performed on the same scale (0, 1, 3, 9). Estimates of the possible reduction r_{ij} of the likelihood and the severity of human error scenario i because of the error blocking by quality system layer j (degree of interaction) are made by the same expert(s) using again the same scale. The interrelationship matrix of r_{ij} has I rows and J columns, as shown in Fig. 2.

Blocking human error according to scenario i by a quality system component j can be more effective in presence of another component j' ($j' \neq j$) because of the synergy $\Delta_{jj'}^{(i)}$ between the two components. The synergy may be equal to 0 or 1 whenever the effect is absent or present, respectively. Estimates q_j of importance/effectiveness of quality system component j in human error reduction are calculated as $q_j = \sum_{i=1}^{I} p_i l_i r_{ij} s_{ij}$, where the synergy factor is $s_{ij} = 1 + \sum_{j' \neq j}^{J} \Delta_{jj'}^{(i)} / (J-1)$. Considering the synergy factor, the interrelationship matrix is to be transformed replacing r_{ij} by $\tilde{r}_{ij} = r_{ij}\, s_{ij}$ in every cell ij of the matrix.

This technique of multi-hazard and multi-risk assessment [12,56,57] allows conversion of the semi-intuitive expert judgments on human errors and on the laboratory quality system into the following quantitative scores expressed in %: a) likelihood score of human error in the analysis $P^* = (100\,\%/9)\sum_{i=1}^{I} p_i/I$; b) severity (loss) score of human error $L^* = (100\,\%/9)\sum_{i=1}^{I} l_i/I$; c) effectiveness score of a component of the laboratory quality system $q_j^* = (100\,\%)\, q_j/\sum_{j=1}^{J} q_j$; and d) effectiveness score of the quality system, as a whole, against human error $E^* = (100\,\%/9)\sum_{j=1}^{J} q_j/\sum_{j=1}^{J}\sum_{i=1}^{I} p_i l_i s_{ij}$. The effectiveness score of the quality system at different steps of the analysis can be evaluated also.

2.4. Evaluation

A score characterizing the risk reduction by the laboratory quality system in whole, expressed in %, is $r^* = (100\,\%/18IJ)\sum_{j=1}^{J}\sum_{i=1}^{I}\tilde{r}_{ij}$ Then, a score of residual risk of human errors (%) which are not prevented/blocked or reduced/mitigated by the quality system, is $R^* = 100\,\% - r^*$ The percentage of the quality of the analytical results which may be lost due to residual risk of human errors is $f_{HE} = (P^*/100\,\%)\,(L^*/100\,\%)\,R^*$

In practice, a quality system is not able to prevent or block human errors completely, i.e. $f_{HE} > 0$ %. Therefore, the residual risk can be interpreted as a source of measurement uncertainty when human being is involved in the measurement process and human interaction with the measuring system is considered. The best metrological situation is when the contribution of the residual risk to the budget of the combined measurement uncertainty is negligible.

As any expert is also a human being [58], the elicitation process of the expert judgements (by which the expert provides error likelihood, severity, and other estimates) is influenced by epistemic uncertainty, intrapersonal conflicts, etc. [59,60] Therefore, evaluation of variability of the error quantification scores and relative risks due to the inherent expert's hesitancy, is important. A detailed analysis of the score variability, as well as variability of corresponding loss of quality f_{HE}, based on Monte Carlo simulations is described in Ref. [32].

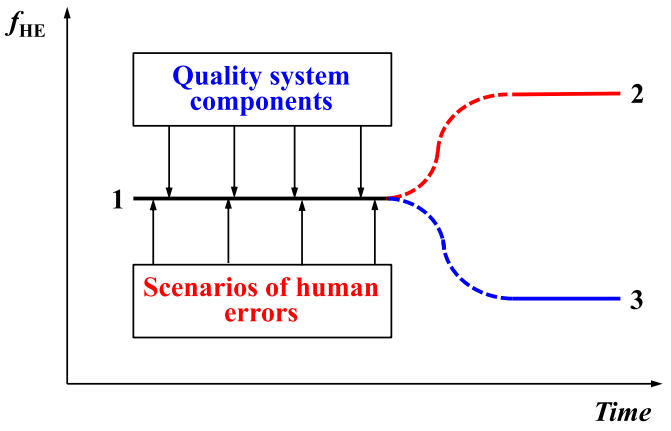


Fig. 3. Quality f_{HE} of analytical results, which may be lost due to residual risk of human errors, vs. time. Estimated current f_{HE} is shown by the flat black line 1; the red curve 2 illustrates possible f_{HE} changes for worse, and the blue curve 3 – for better. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Examples of practical applications of this approach for risk assessment in pH-metric measurements of groundwater, in multi-residue pesticide analysis of fruits and vegetables, and in ICP-MS analysis of geological samples are available in Ref. [32] as well.

2.5. Limitations

Changes in any quality system component over time require re-evaluation of the quality f_{HE} of the analytical results, which may be lost due to residual risk of human errors, estimated currently and shown by the flat black line 1 in Fig. 3.

An f_{HE} increase (for worse) is possible, e.g., due to the retirement of an experienced supervisor, illustrated with the red curve 2. A decrease of f_{HE} depicted by the blue curve 3 may happen (for better), for example, due to the acquisition of a new more accurate and more automated measuring system.

Latent errors due to a poor laboratory design, a defect in the equipment (including a robotic system) or an unsuccessful management decision, as well as positive human factors allowing to overcome the problems, are not considered here.

Another inherent problem of the technique discussed above is that, as a rule, only one expert (supervisor, quality manager or similar) in a laboratory has the necessary knowledge for the elicitation process. His/her estimates are ordinal data, and the obtained dataset is of a small size. Ordinal data can have empirical relations only; they can be equal or unequal, greater than or less than; and should not be treated as continuous quantities [61–63].

Future development of the approach to assessment of risks caused by human errors, based on ordinal data analysis of variation ORDANOVA and a computerized tool [64], could be helpful for analytical chemistry laboratories.

3. Risks of false decisions in conformity assessment due to measurement uncertainty of a chemical analytical result

Conformity assessment of an item (a batch, lot or a sample of a product, material, object, etc.) consists in verifying whether a certain property of the item – the measurand (e.g., i-th component concentration or content c_i under control in the item, $i = 1, 2, \ldots, n$), lies within a specification/tolerance interval $T_i = [T_{Li}, T_{Ui}]$, where T_{Li} and T_{Ui} are the lower and upper tolerance limits of the interval, respectively. These

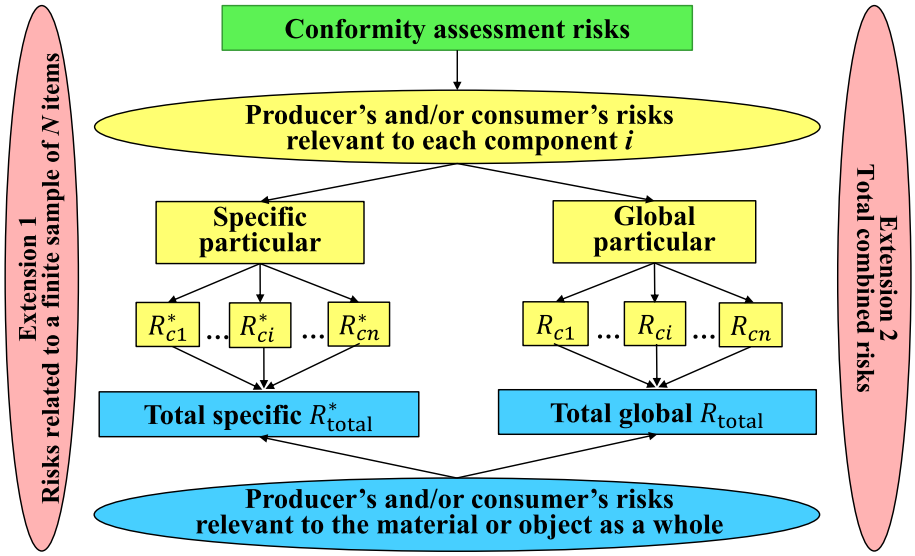


Fig. 4. A map of the risks of false decisions on the conformity of a material or object (green shape). Specific risks refer to a given item, and global risks – to the population of the items. Particular risks (yellow shapes), either specific R^*_{ci} or global R_{ci}, refer to the i-th component of the material under control, $i = 1, \ldots, n$. Total risks (blue shapes), specific R^*_{total} and global R_{total}, are related to the material or object as a whole. These risks are relevant for both the material producer and its consumer. Vertical pink ellipses illustrate extensions to this classification of the risks. Modified from Ref. [37]. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

limits are prescribed in a standard or another regulatory document based on a scientific study, and considered as limits of true property values. However, the true value of the measurand remains unknown since any corresponding measured value has an associated measurement uncertainty [61,63]. Therefore, conformity assessment decisions to accept or reject an item are made using a factory acceptance interval of permissible measured values $A_i = [A_{Li}, A_{Ui}]$, where A_{Li} and A_{Ui} are the lower and upper limits of the interval, respectively. The acceptance interval is usually established considering the factory experience and the measurement uncertainty [65]. It is narrower than the tolerance interval when defending the interests of a consumer, and wider than the tolerance interval when defending the interests of a producer. The difference between corresponding limits of these intervals is named "guard band" [65,66]. The tolerance and acceptance intervals may also coincide [39].

Despite advances in the development of measuring instruments and methodology, to know both the measured value and the acceptance limits with perfect accuracy is impossible. Moreover, comparison of a measured value with the specification/tolerance limits, taken as true values, is still influenced by the measurement uncertainty associated with the measured value. Thus, in general, the uncertainty principle in conformity assessment holds that a decision on conformity or nonconformity (to accept or reject an item) cannot be certain, and risks/probabilities of false/incorrect decisions are never equal to zero [67]. Assessment of such risks in analytical chemistry laboratories described in the guides [36–38] and related publications is discussed further in this review.

3.1. Classification

3.1.1. Particular risks

A component-by-component conformity assessment of an item includes comparison of each measured value c_{im} with the limits of the corresponding acceptance interval A_i, $i = 1, \ldots, n$. The risks expressed as probabilities of false decisions on conformity related to the particular i-th component (of n) are the *particular risks*. Evaluation of such univariate risks is described in JCGM 106 [39] and implemented for analytical chemistry laboratories in Ref. [36].

The probability of a false decision that a component content does not exceed the upper tolerance limit, for example, based on the measured value $c_{im} \le A_{Ui}$, when the true content exceeds the upper tolerance limit ($c_i > T_{Ui}$), is the *particular consumer's risk*. The probability of falsely rejecting conformity ($c_{im} > A_{Ui}$, when in fact $c_i \le T_{Ui}$) is the *particular producer's risk*.

For a specified item, the particular risks related to the i-th component content are referred to as the *particular specific consumer's risk* $R^*_{ci(c)}$ and the *particular specific producer's risk* $R^*_{ci(p)}$, respectively. The particular risks of the incorrect conformity assessment of the i-th component content in an item randomly drawn from a statistical population of such items are the *particular global consumer's risk* $R_{ci(c)}$ and the *particular global producer's risk* $R_{ci(p)}$, respectively.

Note that consequences (severity) of false decisions on conformity of a product depend on numerous factors including the production technology, the purposes of the consumer and the conditions of use of the product. Moreover, there are usually many consumers of a product of one and the same producer. Such consequences of the false decisions are out of a laboratory scope and not considered here.

3.1.2. Total risks

A multicomponent material or object is not simply a mixture of n components or their sum. When the conformity assessment for each i-th component content in an item is successful (*i.e.* the particular risks are small enough), the total probabilities of a false decision concerning the conformity of the item as a whole – the *total risks* – might still be significant. Evaluation of the multivariate total risks is detailed in Ref. [37].

The probability of a false decision that an item is conforming, based on the measured values not exceeding their upper acceptance limits, for example, $c_{im} \le A_{Ui}$ for all $i = 1, \ldots, n$, while as minimum one true component content exceeds its upper tolerance limit ($c_j > T_{Uj}$), is the *total consumer's risk*. If the conformity of a specific item is assessed, this risk is the *total specific consumer's risk* $R^*_{total(c)}$. When an item is randomly drawn from a statistical population of such items, this is the *total global consumer's risk* $R_{total(c)}$.

The probability of falsely rejecting conformity of an item is the *total producer's risk* when, at minimum, one measured value exceeds its upper acceptance limit ($c_{jm} > A_{Uj}$), for example, while all true values satisfy their upper tolerance limits, i.e. $c_i \le T_{Ui}$ for all $i = 1, \ldots, n$. It is the *total specific producer's risk* $R^*_{total(p)}$, when the conformity of a specific item is

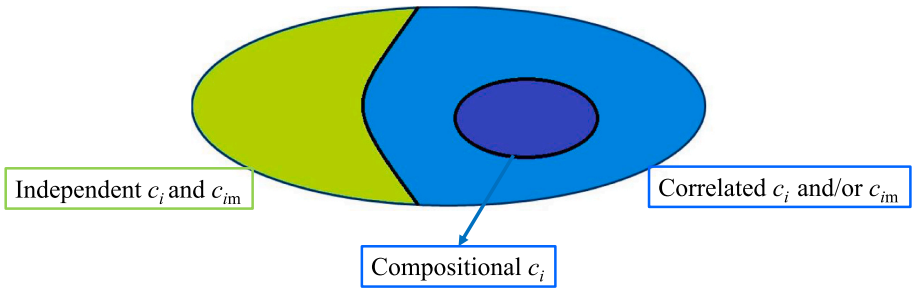


Fig. 5. Venn diagram of data. Independent (hence uncorrelated) actual and measured values are shown as the green part of the data, and correlated values as the light blue part. The dark blue core represents correlated data containing compositional actual values [38]. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

assessed, and the *total global producer's risk* $R_{\text{total(p)}}$, if an item is randomly drawn from a statistical population of such items.

3.1.3. *A map of the risks*

A scheme summarizing the classification of the risks of false decisions on the conformity of a material or object, represented as a map of the relationships among the different kinds of these risks, is shown in Fig. 4. There are four kinds of particular risks for each *i*-th component of the material and four kinds of total risks. Therefore, for $n > 1$ components under control, one can distinguish $4(n+1)$ kinds of risks of false decisions. For example, for two, three, and four components, this means 12, 16, and 20 kinds of risks, respectively: the complexity of conformity assessment increases with the number of components under control [37].

One of the extensions to the classification on the map is related to the risks of false decisions in conformity assessment of a *finite sample of N items* — batches or lots of a material, drug, or food, continuously produced at a factory [67]. Evaluation of such risks is especially important for characterization of a new technological process; when a supplier of raw materials was changed; a key person from the staff and/or a part of the production equipment was replaced; etc. In other words, this is a useful tool for an intermediate analysis of the production success at its installation and any improvement, as well as for monitoring environmental changes.

The second extension, named *total combined risks* was applied to the assessment of the pharmaceutical equivalence of generic and brand-name medicines produced by two corresponding manufacturers. In such a case, the risks of false decisions on the equivalence are a combination of the total risks of conformity assessment of both the products [41,68]. Evaluation of total combined risks can also be applicable when conformity assessment includes testing chemical composition and simultaneously physical, (micro)biological, or sensory characteristics of a substance, material, or object.

3.2. *Modelling*

3.2.1. *Correlations*

Many techniques are used during development and validation of a chemical analytical method to overcome metrological reasons of correlation between measured values c_{im} of the two or more component contents of the same item, $i = 1, 2, \ldots, n$, when their actual (true) values c_i are not correlated. Some of these techniques include extraction of target components (analytes) from a sample and chromatographic separation of an analyte from other components of the sample. Chemometrics software is applied for separation of spectral signals and multivariate calibrations of spectrometers. Sample digestion and standard additions of an analyte to a sample are used for calibration of a measuring system to overcome multiplicative matrix effects, and so on. Correlations that have arisen in the routine measurement process should, in general, be negligible and measured content values c_{im} of two or more components of the same item are expected to be metrologically independent – such data are presented in Fig. 5 as the green part of the Venn diagram.

In practice, if statistically significant correlation between measured values of contents of components of the same item is detected, analysis cannot be continued without a thorough chemical analytical inspection of the reason for the correlation in the laboratory. For example, when a medication is tested routinely with a pharmacopeial HPLC procedure, correlations might be related to the resolution of the chromatography column used, and if it is not able to separate the analytes completely, the column must be replaced by another one. In Ref. [32] such an event is rated as a skill-based mistake or omission error (lapse).

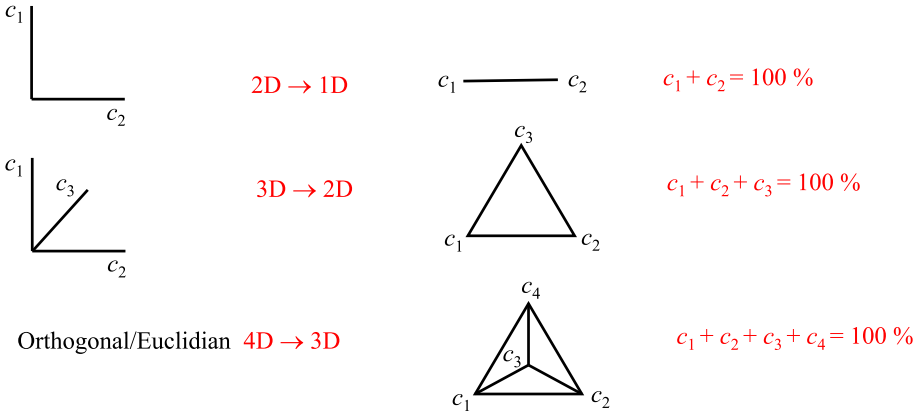


Fig. 6. Orthogonal coordinates of Euclidean spaces (left column) and simplices (right column). The variables c_i (%) are contents of a material or object components $i = 1, 2, 3$, and 4; each simplex vertex corresponds to $c_i = 100$ %; D is dimension. In each case, the mass balance constraint reduces the number of dimensions by one [38].

At the same time, the measured values c_{im} of a particular component i in two or more items (e.g., the active pharmaceutical ingredient in samples of a generic and brand-name medicine) tested in the same laboratory, may not be metrologically independent, since the same equipment and measuring instruments are applied [69].

Moreover, measured values c_{im} of component contents of a substance or material are inevitably correlated when their actual values c_i are correlated. Correlations of true component contents c_i can be caused by natural physico-chemical properties of substances, such as stoichiometry, and by technological reasons in the material production (e.g., in generic reproduction of a brand-name medicine). In any case, the observed correlations are taken into account in conformity assessment using conventional multivariate statistical methods [37], regardless of the correlation origin – metrological, natural, technological or their combination. The correlated data are shown in Fig. 5 as the light blue part of the Venn diagram.

A special case is when component contents of a substance or material are subject to a mass balance constraint, i.e. sum of the actual contents c_i – mass fractions or amount fractions – is 1 or 100 %. Measured values of such component contents are called "compositional data." These data, indicated in Fig. 5 by the dark blue color, are intrinsically correlated and the relevant correlation was named by Karl Pearson in 1897 as "spurious" [38].

3.2.2. Mass balance constraint

There is a strong message in the literature stressing how traditional statistical techniques may produce inadequate results if applied to raw compositional data [70]. Mathematical aspects of Compositional Data Analysis (CoDA) based on an isometric logratio transformation of the original measured values and its applications in environmental analysis, geology and other fields have been intensively developed [71–73]. However, CoDA applies the constraint directly to the measured values, without considering their associated measurement uncertainties.

When n true (actual) component contents c_i of a substance or material are subject to a mass balance constraint ($\sum_{i=1}^{n} c_i = 1$ or 100 %), the component contents are compositional data, that may be depicted in a multidimensional simplex, as in Fig. 6, in which Euclidean geometry cannot, in general, be applied blindly. For the model $c_1 + c_2 = 100$ %, the correlation between the two component contents is exactly equal to −1. For the model $c_1 + c_2 + c_3 = 100$ %, the first component content can be derived from the other two, i.e., $c_1 = 100\ \% - c_2 - c_3$, and the corresponding Pearson correlation coefficients are $(r_{12})_{an} = -s_2^2/(s_2^2+s_3^2)$ and $(r_{13})_{an} = -s_3^2/(s_2^2+s_3^2)$, respectively, where s_i is the standard deviation of the i-th component content c_i actual value. Note that a mass balance constraint is applicable to actual component contents c_i, while the sum of measured values c_{im} can, in general, be different from 1 (or 100 %, or another constant) because of measurement uncertainties u_{im} associated with the measured values.

The vector of actual component content values $\boldsymbol{c} = [c_1, c_2, \ldots, c_n]$ describes the n-component (n-part) composition of a material or substance and lies within the compositional space, i.e., the simplex

$$S^n = \left\{ \boldsymbol{c} = [c_1, c_2, \ldots, c_n] \;\middle|\; c_i > 0, i = 1, 2, \ldots, n;\ \sum_{i=1}^{n} c_i = k \right\}, \tag{1}$$

where k is usually equal to 1 or 100 %. As c_i are positive quantity ratios, a vector $\boldsymbol{c}$ multiplied by any positive constant retains the same information as the original one, i.e., represents the same composition and can be considered as an equivalence class. This property is termed "scale invariance." In other words, if $\boldsymbol{c}$ is scaled by a constant, e.g., content values c_i change from parts-per-unit to percentages, the information which $\boldsymbol{c}$ conveys is completely equivalent. Therefore, it is natural to select a representative of the equivalence class to facilitate data analysis and interpretation of corresponding results. This selection is formalized by the closure operation:

$$clo\ (\boldsymbol{c}) = \left[\frac{k \cdot c_1}{\sum_{i=1}^{n} c_i}, \ldots, \frac{k \cdot c_n}{\sum_{i=1}^{n} c_i} \right]. \tag{2}$$

As vector $\boldsymbol{c}$ represents the actual n-component composition of a substance or material, component contents c_i are the measurands in conformity assessment, and $\boldsymbol{c}$ is the vector of the measurands.

3.2.3. Bayesian approach

According to Bayes' theorem, a prior knowledge of the measurand values (actual contents $\boldsymbol{c}$ of components in the items) and new information acquired during the measurement (modelled by a likelihood function) are combined in a posterior probability density function (pdf) [74]. Such posterior pdf, containing all the available information, is then used for the evaluation of the risks [8,39].

Knowledge about a composition $\boldsymbol{c}$ can be modelled by a random multivariate posterior variable and expressed in terms of its probability density function (pdf). Such a pdf combines prior knowledge about the measurands and new information acquired during the measurements:

$$g(\boldsymbol{c} \mid \boldsymbol{c}_m) = C g_0(\boldsymbol{c}) h(\boldsymbol{c}_m \mid \boldsymbol{c}), \tag{3}$$

where $\boldsymbol{c}_m$ is the vector of measured content values c_{im}; $g(\boldsymbol{c} \mid \boldsymbol{c}_m)$ is the multivariate posterior pdf, i.e. the probability density function of actual values $\boldsymbol{c}$ when the obtained measured values are $\boldsymbol{c}_m$; C is a normalizing constant; $g_0(\boldsymbol{c})$ is the multivariate prior pdf; and $h(\boldsymbol{c}_m \mid \boldsymbol{c})$ is the multivariate likelihood function taking into account the measurement uncertainties and possible covariance terms, characterizing the distribution of measured values $\boldsymbol{c}_m$ for one and the same sample of the item with actual component content values $\boldsymbol{c}$.

The total global risks of a false decision on the conformity of a multicomponent material or object are evaluated according to the following formulae, for the consumer and the producer, respectively:

$$R_c = \int_{T^c} \int_{A} g_0(\boldsymbol{c}) h(\boldsymbol{c}_m \mid \boldsymbol{c})\ \mathrm{d}\boldsymbol{c}_m \mathrm{d}\boldsymbol{c}, \text{ and} \tag{4}$$

$$R_p = \int_{T} \int_{A^c} g_0(\boldsymbol{c}) h(\boldsymbol{c}_m \mid \boldsymbol{c})\ \mathrm{d}\boldsymbol{c}_m \mathrm{d}\boldsymbol{c}, \tag{5}$$

where T is the multivariate tolerance/specification domain $T_1 \times T_2 \times \cdots \times T_n$; A is the multivariate acceptance domain $A_1 \times A_2 \times \cdots \times A_n$; the integral symbols indicate multiple integrals; superscript "c" of T in the formula for R_c means "complementary" for at least one T_i, whereas the integration with respect to all c_{im} is performed within A; the superscript "c" of A in the formula for R_p means "complementary" for at least one A_i, whereas the integration with respect to all c_i is performed within T.

The total specific risks for the consumer and the producer are calculated, respectively, by:

$$R_c^* = 1 - \int_{T} g(\boldsymbol{c} \mid \boldsymbol{c}_m) \mathrm{d}\boldsymbol{c} \text{ when } \boldsymbol{c}_m \text{ is in } A, \text{ and} \tag{6}$$

$$R_p^* = \int_{T_1} \ldots \int_{T_\nu} \int_{0}^{100} \ldots \int_{0}^{100} g(\boldsymbol{c} \mid \boldsymbol{c}_m) \mathrm{d}\boldsymbol{c} \text{ when } c_{im} \text{ are outside } A \text{ for } 1 \le i \le \nu. \tag{7}$$

Here R_c^* is the probability that at least one of corresponding actual content values c_i is outside its tolerance interval, when all the measured content values c_{im} are within their acceptance intervals (false conforming); symbol ν in Eq. (7) for R_p^* indicates the number of those components whose measured content values c_{im} are outside their acceptance intervals, $\nu \le n$.

3.2.4. Prior pdf

When a large enough dataset of results of tested items of the same material produced at the same factory is available in the laboratory, it can be used for approximation of the prior pdf. This is also possible with

the corresponding dataset of results of monitoring the same environmental compartment. The assumption is that the actual content values c_i are approximated by the test/measurement results c_{im} adequately, when measurement uncertainty is negligible in comparison with item-to-item (batch-to-batch) variations caused by intrinsic variability of conditions of the material production, environmental conditions, etc.

For modelling univariate risks in conformity assessment of a single component, the prior pdfs in Ref. [36] were approximated with lognormal distributions for contents of total suspended particulate matter (TSPM) in ambient air of industrial zones; the Weibull distribution was applied for contents of pesticide residuals in tomatoes; Student's and normal distributions were used for the contents of active pharmaceutical ingredients (APIs) in some drugs. Multivariate risks were modelled in Ref. [37] with the prior pdfs of the independent contents of three denaturants in alcohol for industrial use, approximated by normal distributions; TSPM contents in ambient air around three different quarries – by lognormal distributions (no correlation among test results was observed); correlated contents of four APIs in a cold/flu medication, and correlated contents of four components of a platinum-rhodium (PtRh) alloy – by multivariate normal distributions.

For a case of mass balance constraint, the data properties need to be taken into account when assigning a corresponding prior pdf [38]. As the vector $\boldsymbol{c}$ is non-negative, and the vector of mean values $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{V}$ are only assumed to be known, a truncated multivariate normal distribution TMN($\boldsymbol{\mu}$,$\boldsymbol{V}$) on the nD region $[0, k]^n$ corresponds to the principle of maximum entropy distribution [75,76]. Therefore, it was applied for the prior pdf modelling. Three scenarios are considered:

1) modelling all the actual values of the components' contents by applying the closure operation $clo(\boldsymbol{c})$ to $\boldsymbol{c}$ which follows a TMN($\boldsymbol{\mu}$,$\boldsymbol{V}$);
2) modelling actual values of n-1 components' contents by a TMN($\boldsymbol{\mu}'$,$\boldsymbol{V}'$) on a (n-1)D region and then deriving the n-th as 100 % minus the sum of the other components' values; and
3) sequentially modelling one component at a time by a (univariate) truncated normal distribution, whose domain of definition depends on the previously modelled component c_i value.

These modelling approaches were applied in Ref. [38] for evaluation of risks in conformity assessment of the PtRh alloy, pure potassium trioxidoiodate, a sausage, and synthetic air. When a suitable experimental dataset is absent, selection of an appropriate prior distribution in risk assessment can be based on the theoretical knowledge about physico-chemical and other properties of the item under study, and common sense [77,78].

3.2.5. Likelihood function

A distribution of measured values $\boldsymbol{c}_\mathbf{m}$ at a given actual content values $\boldsymbol{c}$ of components of a multicomponent material or object, is caused by measurement uncertainties u_i (in general, by covariance matrix $\boldsymbol{U}$ associated with $\boldsymbol{c}_\mathbf{m}$), that are available from the analytical method validation data. Note that the contribution of measurement uncertainty, arising from sampling a test item, may be important in conformity assessment [79,80].

For modelling univariate risks of a single component and multivariate risks in conformity assessment of a multicomponent material or object, mentioned in Sec. 3.2.4, normal distributions were applied.

The modelling of the likelihood function for measured content values $\boldsymbol{c}_\mathbf{m}$ in the case of a mass balance constraint is based on a choice of an appropriate pdf with zero expectation for an error vector $\boldsymbol{e}_\mathbf{m}$, translated then to the vector of actual content values $\boldsymbol{c}$ generated according to the prior. Thus, vector $\boldsymbol{c}_\mathbf{m}$ is recovered as $\boldsymbol{c}_\mathbf{m} = \boldsymbol{c} + \boldsymbol{e}_\mathbf{m}$. The covariance matrix $\boldsymbol{U}$ associated with $\boldsymbol{c}_\mathbf{m}$ contains the squared measurement uncertainties u_i^2 and the covariance terms u_{ij} whose corresponding correlation coefficients are the same as for the covariance matrix $\boldsymbol{V}$ associated with $\boldsymbol{c}$. The modelling of the likelihood function for such cases follows that of the prior by the three scenarios above in Sec. 3.2.4.

3.2.6. Posterior pdf

The posterior distribution is the normalized product of the prior and the likelihood according to Eq. (3). Since the posterior pdf predicts further actual content values (after accumulation of the dataset used for modelling the prior pdf), taking into account what may happen during the measurement process *via* the likelihood function, the closure operation is not appropriate for the posterior data. In other words, the sum of the calculated (predicted) actual component content values may differ from 100 % because any predicted value has its associated uncertainty.

3.3. Evaluation

When prior pdfs and likelihood functions are modelled using analytical expressions for the approximating distributions as in Refs. [36,37], the risks may be evaluated just by Eqs. (4)–(7), either using analytical calculations or numerically.

For evaluation of the total global risks in the case of variables related to a mass balance, a Monte Carlo numerical solution is typically the most helpful [38]. The total global producer's risk is evaluated considering the fraction of simulated vectors in which all the actual values are within the corresponding tolerance region T, while at least one of the measured values is out of its acceptance interval A_i. The total global consumer's risk is estimated as the fraction of vectors in which all the measured values are within the acceptance region A, while at least one of the actual values is out of its tolerance interval T_i.

As for the total specific risks, for each specified vector of measured values, the integrals of the posterior pdf involve the ratio of multiple integrals of the joint pdf with respect to variables c_i over appropriate domains. The numerical evaluation of such integrals is performed again by simulation of M random vectors $[c_1, \ldots, c_n, c_{1m}, \ldots, c_{nm}]$, generated according to the prior modelling for c_i values and the likelihood modelling for c_{im} values.

An inverse task is recalculation of the acceptance limits (region A) using the known tolerance limits and obtained estimates of the total risks for reduction of the risks to admissible values [81,82]. Note that in practice, a change of acceptance limits reducing the risks simultaneously for both, consumer and producer, is possible only by decreasing the measurement uncertainties u_i through investment in new equipment and measuring instruments. Decreasing item-to-item (batch-to-batch) variations, i.e. the terms of the covariance matrix $\boldsymbol{V}$, for reducing the risks is out of the laboratory scope.

Available examples of practical applications of this approach for assessment of the risks of false decisions, mentioned already in Sec. 3.2.4, are related to investigations of out-of-specification test results of concentrations of total suspended particulate matter in ambient air of industrial zones; pesticide residues in tomatoes; sodium chloride and L-adrenaline in a long-term study of drugs; and cetirizine dihydrochloride assay in a bulk material [36]. Examples of assessment of risks of false decisions on conformity of multicomponent materials were developed for custom control of denatured alcohols; monitoring total suspended particulate matter in ambient air contributed by three independent stone quarries; control of composition of a cold/flu medication; and of a PtRh alloy [37]. Mass balance constraints were considered for assessment of risks of false decisions on conformity of the PtRh alloy evaluated in the previous study; conformity of pure KIO_3; a summer (dry) sausage; and synthetic air [38].

3.4. Limitations

As per usual, limitations apply to the modelling and evaluation:

a) the use of any model as a simplified reflection of reality;
b) adequacy of the assumption of negligible definitional uncertainty of actual component contents c_i including inhomogeneity and/or instability of an item of the multicomponent substance, material, or object;

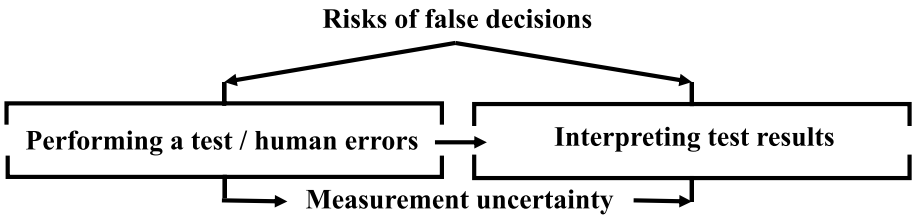


Fig. 7. A scheme of risks of false decisions in an analytical chemistry laboratory. Risks in performing a test, caused by human errors not reduced by the quality system, affect measurement uncertainty which is the reason for false decisions in interpreting test results at conformity assessment.

c) adequacy of the treatment of a dataset of item-to-item (batch-to-batch) test/measurement results for modelling a prior pdf;
d) goodness-of-fit of experimental and theoretical distributions, etc.
e) limitations of applied computational methods (e.g., Monte Carlo) and programming.

It is expected that all these limitations (as well as those in modelling and evaluation of risks caused by human errors) will be minimized in the near future with the integration of artificial intelligence (AI) in analytical chemistry (testing) laboratories for risk assessment [83].

4. Conclusions

There are two groups of decision-making risks in an analytical chemistry (testing) laboratory that directly influence quality of measurement/test results, summarized schematically in Fig. 7.

The first group consists of the risks of false decisions caused by human errors in performing a test. The approach to modelling and evaluation of such risks is based on estimates/judgements of a competent laboratory expert (supervisor, quality manager or similar) who has the necessary knowledge for the elicitation process using an ordinal scale. The applied mathematics is simple. However, a laboratory ordinal dataset is of a small size, and the obtained estimates change with changes in the laboratory staff and/or equipment. Future development of this approach using ordinal data analysis of variation ORDANOVA and a computerized tool, could be helpful.

The second group of risks is from false decisions in conformity assessment of test results judged against the specification/tolerance limits due to measurement uncertainty. These risks are modelled and evaluated using a large enough dataset of the laboratory for different batches or lots of the product (or results of monitoring an environment compartment), by the Bayesian approach, and multivariate statistics that are able to overcome correlations of the data and a mass balance constraint.

It is expected that limitations of the modelling and evaluation of risks of both groups will be minimized with integration of AI in analytical chemistry (testing) laboratories for risk assessment.

CRediT authorship contribution statement

Ilya Kuselman: Writing – original draft, Visualization, Investigation, Conceptualization. **Francesca R. Pennecchi:** Writing – review & editing, Software, Methodology, Formal analysis. **D. Brynn Hibbert:** Writing – review & editing, Supervision. **Angelique Botha:** Writing – review & editing, Validation, Data curation. **Tamar Gadrich:** Writing – review & editing, Validation. **Anastasia A. Semenova:** Writing – review & editing, Validation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We acknowledge organizational support from the International Union of Pure and Applied Chemistry, IUPAC Project 2024-012-2-500.

Data availability

Data will be made available on request.

References

[1] V.T. Covello, J. Mumpower, Risk analysis and risk management: an historical perspective, Risk Anal. 5 (1985), https://doi.org/10.1111/j.1539-6924.1985.tb00159.x.
[2] Christian Bible, *Proverbs 22:3*, New Living Translation. https://www.bible.com/bible/116/PRO.22.3.NLT. (Accessed 11 November 2024).
[3] H.F. Kloman, A brief history of risk management, in: Enterprise Risk Management, 2009, pp. 19–29, https://doi.org/10.1002/9781118267080.ch2.
[4] Y. Li, T. Hills, R. Hertwig, A brief history of risk, Cognition 203 (2020), https://doi.org/10.1016/j.cognition.2020.104344.
[5] T. Aven, Risk assessment and risk management: review of recent advances on their foundation, Eur. J. Oper. Res. 253 (2016), https://doi.org/10.1016/j.ejor.2015.12.023.
[6] International Organization for Standardization ISO 31073:2022, Risk Management — Vocabulary, ISO, Geneva, Switzerland, 2022. https://www.iso.org/standard/79637.html.
[7] International Organization for Standardization ISO 31000:2018, Risk Management — Guidelines, ISO, Geneva, Switzerland, 2018. https://www.iso.org/standard/65694.html.
[8] International Organization for Standardization IEC 31010:2019, Risk Management — Risk Assessment Techniques, ISO, Geneva, Switzerland, 2019. https://www.iso.org/standard/72140.html.
[9] Risk management, ISO/TC 262. https://committee.iso.org/home/tc262, Geneva, Switzerland, Accessed: 11 November 2024. .
[10] International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH), Q9(R1) Quality Risk Management, Guidance for Industry, U.S. Department of Health and Human Services Food and Drug Administration, New Hampshire, USA, 2023. https://www.fda.gov/media/167721/download.
[11] International Organization for Standardization ISO 22367:2020 Medical Laboratories — Application of Risk Management to Medical Laboratories, ISO, Geneva, Switzerland, 2020. https://www.iso.org/standard/71254.html.
[12] European Commission Joint Research Centre, Science for Disaster Risk Management 2020 – Acting Today, Protecting Tomorrow, Publications Office, 2021, https://doi.org/10.2760/438998.
[13] K.R. Solomon, A brief history of risk assessment for agrochemicals, Pure Appl. Chem. 96 (2024) 1643, https://doi.org/10.1515/pac-2024-0215.
[14] S. A. European Commission, Directorate-general for employment, inclusion. Guidance on Risk Assessment at Work, Publications Office, 1996. https://op.europa.eu/en/publication-detail/-/publication/1a3462b0-728c-4a2b-88f0-6c641b91a86f.
[15] V. Karahan, E. Aydoğmuş, Risk analysis and risk assessment of laboratory work by fine kinney method, Int. J. Adv. Nat. Sci. Eng. Res. 7 (2023) 442, https://doi.org/10.59287/ijanser.788.
[16] E. Tziakou, A.G. Fragkaki, A.N. Platis, Identifying risk management challenges in laboratories, Accred Qual. Assur. 28 (2023), https://doi.org/10.1007/s00769-023-01540-3.
[17] Z. Li, X. Wang, S. Gong, N. Sun, R. Tong, Risk assessment of unsafe behavior in university laboratories using the HFACS-UL and a fuzzy Bayesian network, J. Saf. Res. 82 (2022) 13, https://doi.org/10.1016/j.jsr.2022.04.002.
[18] R.W. McLeod, Approaches to understanding human behavior when investigating incidents in academic chemical laboratories, ACS Chem. Health Saf. 29 (2022) 263, https://doi.org/10.1021/acs.chas.2c00020.
[19] G. Topal, S. Şanlı, Risk assessment in a public university chemistry laboratories, J. Int. Health Sci. Manag. 7 (2021) 17, https://doi.org/10.48121/jihsam.892548.
[20] H. Abedsoltan, M.B. Shiflett, Mitigation of potential risks in chemical laboratories: a focused review, ACS Chem. Health Saf. 31 (2024) 104, https://doi.org/10.1021/acs.chas.3c00097.

[21] J. Zhao, H. Cui, G. Wang, J. Zhang, R. Yang, Risk assessment of safety level in university laboratories using questionnaire and Bayesian network, J. Loss Prev. Process. Ind. 83 (2023) 105054, https://doi.org/10.1016/j.jlp.2023.105054.
[22] C. Xu, L. Guo, K. Wang, T. Yang, Y. Feng, H. Wang, D. Li, G. Fu, Current challenges of university laboratory: characteristics of human factors and safety management system deficiencies based on accident statistics, J. Saf. Res. 86 (2023), https://doi.org/10.1016/j.jsr.2023.07.010.
[23] C. Bellini, R. Guerranti, F. Cinci, E. Milletti, C. Scapellato, Defining and managing the preanalytical phase with FMECA: automation and/or "Human" control, Hum. Factors 62 (2020) 20, https://doi.org/10.1177/0018720819874906.
[24] J. West, J. Atherton, S.J. Costelloe, G. Pourmahram, A. Stretton, M. Cornes, Preanalytical errors in medical laboratories: a review of the available methodologies of data collection and analysis, Ann. Clin. Biochem. 54 (2017) 14, https://doi.org/10.1177/0004563216669384.
[25] E. Lao, Á. García, M.M. Figuerola, E, A. Paraire, Errors of clinical laboratory and its impact on patient safety, Open J. Soc. Sci. 5 (2017) 243, https://doi.org/10.4236/jss.2017.53022.
[26] S.A. Farzam, Z. Sahrael, E. Khodabandehloo, H.S. Rad, M.R. Modabber, Y. K. Monfared, Evaluation of human errors using standardized plan analysis risk among health provider in clinical and pathology laboratories in hospitals of Qazvin province, Helix 9 (2019) 5084. https://helixscientific.pub/index.php/home/article/view/1.
[27] M. Vogeser, C. Seger, Irregular analytical errors in diagnostic testing – a novel concept, Clin. Chem. Lab. Med. 56 (2018) 386, https://doi.org/10.1515/cclm-2017-0454.
[28] A.S. Chaudhry, Y. Inata, E. Nakagami-Yamaguchi, Quality analysis of the clinical laboratory literature and its effectiveness on clinical quality improvement: a systematic review, J. Clin. Biochem. Nutr. 73 (2023) 108, https://doi.org/10.3164/jcbn.23-22.
[29] M.M.O. Alrashidi, A.S.W. Alharbi, A.S.W. Alharbi, M.S.A. Alharbi, M.K. M. Almutairi, N.M.A. Alharbi, Analyzing errors in laboratory testing: causes and prevention strategies, Migr. Lett. 19 (2022) 1552. https://migrationletters.com/index.php/ml/article/view/10238.
[30] International Organization for Standardization ISO 15189:2022 Medical Laboratories — Requirements for Quality and Competence, ISO, Geneva, Switzerland, 2022. https://www.iso.org/standard/76677.html.
[31] International Organization for Standardization ISO/IEC 17025, General Requirements for the Competence of Calibration and Testing Laboratories, ISO/IEC, Geneva, 2017, 2017, https://www.iso.org/ISO-IEC-17025-testing-and-calibration-laboratories.html.
[32] I. Kuselman, F. Pennecchi, IUPAC/CITAC guide: classification, modeling and quantification of human errors in a chemical analytical laboratory (IUPAC technical report), Pure Appl. Chem. 88 (2016) 477, https://doi.org/10.1515/pac-2015-1101.
[33] A.B. Bijelić, B.D. Bijelić, D.G. Stanković, D.L. Adamović, T.D. Golubović, S. M. Glišović, E.I. Jovanović, Assessment of human errors in the determination of the concentration of water pollutants, Pol. J. Environ. Stud. (2024), https://doi.org/10.15244/pjoes/188034.
[34] T. Badrick, A.S. Brown, Identifying human factors as a source of error in laboratory quality control, J. Lab. Precis. Med. 8 (2023), https://doi.org/10.21037/jlpm-23-7.
[35] Ya V. Tarasenko, O.V. Chervotoka, M.O. Gerashchenko, I.I. Zaiets, A. M. Gordiienko, Assessment of the influence of human factor on the results of laboratory measurements of weapons and military samples, Proceedings of the State Research Institute of Testing and Certification of Weapons and Military Equipment 20 (2024) 110, https://doi.org/10.37701/dndivsovt.20.2024.15.
[36] I. Kuselman, F. Pennecchi, C. Burns, A. Fajgelj, P. de Zorzi, IUPAC/CITAC guide: investigating out-of-specification test results of chemical composition based on metrological concepts (IUPAC technical report), Pure Appl. Chem. 84 (2012) 1939, https://doi.org/10.1351/PAC-REP-11-10-04.
[37] I. Kuselman, F.R. Pennecchi, R.J.N.B. da Silva, D.B. Hibbert, IUPAC/CITAC guide: evaluation of risks of false decisions in conformity assessment of a multicomponent material or object due to measurement uncertainty (IUPAC technical report), Pure Appl. Chem. 93 (2021) 113, https://doi.org/10.1515/pac-2019-0906.
[38] F.R. Pennecchi, I. Kuselman, D.B. Hibbert, IUPAC/CITAC guide: evaluation of risks of false decisions in conformity assessment of a substance or material with a mass balance constraint (IUPAC technical report), Pure Appl. Chem. 95 (2023) 1217, https://doi.org/10.1515/pac-2022-0801.
[39] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, and OIML JCGM 106:2012, Evaluation of measurement data – the role of measurement uncertainty in conformity assessment, Joint Committ. Guides in Metrol. (2012). Sèvres, France. https://www.bipm.org/documents/20126/2071204/JCGM_106_2012_E.pdf/fe9537d2-e7d7-e146-5abb-2649c3450b25.
[40] L. Separovic, R.S. Simabukuro, A.R. Couto, M.L.G. Bertanha, F.R.S. Dias, A.Y. Sano, A.M. Caffaro, F.R. Lourenço, Measurement uncertainty and conformity assessment applied to drug and medicine analyses – a review, Crit. Rev. Anal. Chem. 53 (2023), https://doi.org/10.1080/10408347.2021.1940086.
[41] A.M. Caffaro, F.R. Lourenço, Total combined global risk assessment applied to pharmaceutical equivalence – a case study of ofloxacin medicines, Chemometr. Intell. Lab. Syst. 241 (2023), https://doi.org/10.1016/j.chemolab.2023.104935.
[42] G.K.T.P. do Amaral, F.R. Lourenço, Risk of false conformity decisions due to measurement uncertainty of active pharmaceutical ingredient: a multiparametric evaluation, Braz. J. Pharm. Sci. 60 (2024), https://doi.org/10.1590/s2175-97902024e23565.
[43] I.E. Dror, Cognitive and human factors in expert decision making: six fallacies and the eight sources of bias, Anal. Chem. 92 (2020), https://doi.org/10.1021/acs.analchem.0c00704.
[44] I.E. Dror, The most consistent finding in forensic science is inconsistency, J. Forensic Sci. 68 (2023), https://doi.org/10.1111/1556-4029.15369.
[45] J. Kukucka, I.E. Dror, in: D. DeMatteo, K.C. Scherr (Eds.), The Oxford Handbook of Psychology and Law, Oxford University Press, Oxford, UK, 2023, pp. 621–642, https://doi.org/10.1093/oxfordhb/9780197649138.013.36.
[46] D. Krewski, P. Saunders-Hastings, P. Larkin, M. Westphal, M.G. Tyshenko, W. Leiss, M. Dusseault, M. Jerrett, D. Coyle, Principles of risk decision-making, J. Toxicol. Environ. Health, Part B 25 (2022), https://doi.org/10.1080/10937404.2022.2107591.
[47] P.J. Borman, A.M. Guiraldelli, J. Weitzel, S. Thompson, J. Ermer, J.-M. Roussel, J. Marach, S. Sproule, H.N. Pappa, Ongoing analytical procedure performance verification using a risk-based approach to determine performance monitoring requirements, Anal. Chem. 96 (2024), https://doi.org/10.1021/acs.analchem.3c03708.
[48] IUPAC Project 2024-012-2-500. Advanced methods for assessment of risks of false decisions in analytical chemistry (testing) laboratories – basic concepts and associated terms, International Union of Pure and Applied Chemistry. https://iupac.org/project/2024-012-2-500, Accessed: 11 November 2024.
[49] Y. Zhao, A Bayesian approach to comparing human reliability analysis methods using human performance data, Reliab. Eng. Syst. Saf. 219 (2022) 108213, https://doi.org/10.1016/j.ress.2021.108213.
[50] L.-X. Hou, R. Liu, H.-C. Liu, S. Jiang, Two decades on human reliability analysis: a bibliometric analysis and literature review, Ann. Nucl. Energy 151 (2021) 107969, https://doi.org/10.1016/j.anucene.2020.107969.
[51] B. Kirwan, in: A Guide to Practical Human Reliability Assessment, first ed., 2017, https://doi.org/10.1201/9781315136349. London.
[52] I. Kuselman, F. Pennecchi, W. Bich, D. Brynn Hibbert, Human being as a part of measuring system influencing measurement results, Accred Qual. Assur. 21 (2016) 421, https://doi.org/10.1007/s00769-016-1239-3.
[53] Center for Drug Evaluation and Research (CDER), Investigating out-of-specification (OOS) test results for pharmaceutical production - level 2 revision: Guidance for industry, U.S. Department of Health and Human Services Food and Drug Administration (2022). https://www.fda.gov/regulatory-information/search-fda-guidance-documents/investigating-out-specification-oos-test-results-pharmaceutical-production-level-2-revision.
[54] J. Reason, Human error: models and management, BMJ 320 (2000) 768, https://doi.org/10.1136/bmj.320.7237.768.
[55] S. Dror, E. Bashkansky, R. Ravid, House of security: a structured system design & analysis approach, Int. J. of Safety and Security Eng. 2 (2012) 317. https://www.witpress.com/elibrary/sse-volumes/2/4/668.
[56] C.J. White, M.S.G. Adnan, M. Arosio, S. Buller, Y. Cha, R. Ciurean, J.M. Crummy, M. Duncan, J. Gill, C. Kennedy, E. Nobile, L. Smale, P.J. Ward, Review article: Towards multi-hazard and multi-risk indicators – a review and recommendations for development and implementation, Nat. Hazards Earth Syst. Sci. Discuss. 2024 (2024) 1, https://doi.org/10.5194/nhess-2024-178.
[57] S. Hochrainer-Stigler, R. Šakić Trogrlić, K. Reiter, P.J. Ward, M.C. de Ruiter, M. J. Duncan, S. Torresan, R. Ciurean, J. Mysiak, D. Stuparu, S. Gottardo, Toward a framework for systemic multi-hazard and multi-risk assessment and management, iScience 26 (2023) 106736, https://doi.org/10.1016/j.isci.2023.106736.
[58] K. Smith, An Expert – Common Key Characteristics (2021) 1, https://doi.org/10.20935/AL4181. Academia Letters https://doi.org/10.20935/AL4181.
[59] R.M. Cooke, L.H.J. Goossens, Expert judgement elicitation for risk assessments of critical infrastructures, J. Risk Res. 7 (2004) 643, https://doi.org/10.1080/1366987042000192237.
[60] A. O'Hagan, Eliciting and using expert knowledge in metrology, Metrologia 51 (2014) S237, https://doi.org/10.1088/0026-1394/51/4/S237.
[61] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, OIML JCGM 200:2012. *International Vocabulary of Metrology – Basic and General Concepts and Associated Terms (VIM)*, 2012. Sèvres, France.
[62] International Organization for Standardization ISO 3534-2:2006, Statistics – Vocabulary and Symbols – Part 2: Applied Statistics, ISO, Geneva, Switzerland, 2006.
[63] D.B. Hibbert (Ed.), Compendium of Terminology in Analytical Chemistry: IUPAC Orange Book, fourth ed., RSC, London, 2023 https://doi.org/10.1039/9781788012881.
[64] T. Gadrich, Y.N. Marmor, F.R. Pennecchi, D.B. Hibbert, A.A. Semenova, I. Kuselman, Power of a test for assessing interlaboratory consensus of nominal and ordinal characteristics of a substance, material, or object, Metrologia 61 (2024) 045004, https://doi.org/10.1088/1681-7575/ad5846.
[65] A. Williams, B. Magnusson (Eds.), Eurachem/CITAC Guide: Use of Uncertainty Information in Compliance Assessment, second ed., 2021. Available from: www.eurachem.org.
[66] ILAC ILAC-G8, Guidelines on decision rules and statements of conformity, 2019. Available from: www.ilac.org, 2019.
[67] F. Pennecchi, I. Kuselman, Probabilities of true and false decisions in conformity assessment of a finite sample of items, Meas. Sci. Technol. 35 (2024) 056003, https://doi.org/10.1088/1361-6501/ad2439.
[68] M.L. de Godoy Bertanha, F.R. Lourenço, Total combined risk of false conformity decisions in pharmaceutical equivalence of acetaminophen oral solution, Microchem. J. 201 (2024) 110610, https://doi.org/10.1016/j.microc.2024.110610.
[69] M.L. de Godoy Bertanha, F.R. Lourenço, Impact of metrological correlation on the total combined risk in pharmaceutical equivalence evaluations, Chemometr. Intell. Lab. Syst. 255 (2024) 105267, https://doi.org/10.1016/j.chemolab.2024.105267.
[70] G. Coenders, J.J. Egozcue, K. Fačevicová, C. Navarro-López, J. Palarea-Albaladejo, V. Pawlowsky-Glahn, R. Tolosana-Delgado, 40 years after Aitchison's article "The

statistical analysis of compositional data". Where we are and where we are heading, SORT 47 (2023) 207, https://doi.org/10.57645/20.8080.02.6.
[71] M. Greenacre, Compositional data analysis, Annual Review of Statistics and Its Application 8 (2021) 271, https://doi.org/10.1146/annurev-statistics-042720-124436.
[72] Peter Filzmoser, Karel Hron, Josep Antoni Martín-Fernández, J. Palarea-Albaladejo, Advances in compositional data analysis, Switzerland,, https://doi.org/10.1007/978-3-030-71175-7, 2021.
[73] A.R. Bennett, J. Lundstrøm, S. Chatterjee, M. Thaysen-Andersen, D. Bojar, Compositional data analysis enables statistical rigor in comparative glycomics, Nat. Commun. 16 (2025) 795, https://doi.org/10.1038/s41467-025-56249-3.
[74] J. Meija, O. Bodnar, A. Possolo, Ode to Bayesian methods in metrology, Metrologia 60 (2023), https://doi.org/10.1088/1681-7575/acf66b.
[75] A. Taavitsainen, R. Vanhanen, On the maximum entropy distributions of inherently positive nuclear data, Nucl. Instrum. Methods Phys. Res. Sect. A Accel. Spectrom. Detect. Assoc. Equip. 854 (2017), https://doi.org/10.1016/j.nima.2016.11.061.
[76] L. Wang, L. Wei, J. Liu, F. Qian, A probability distribution model based on the maximum entropy principle for wind power fluctuation estimation, Energy Rep. 8 (2022), https://doi.org/10.1016/j.egyr.2022.03.186.
[77] D. Bozic, B. Runje, Selection of an Appropriate Prior Distribution in Risk Assessment in *33rd DAAAM International Symposium*, DAAAM International, Vienna, Austria, 2022, pp. 471–479, https://doi.org/10.2507/33rd.daaam.proceedings.066.
[78] D.B. Hibbert, Bayesian approaches to assigning the source of an odour detected by an electronic nose, Aust. J. Chem. 77 (2024), https://doi.org/10.1071/CH24053.
[79] M.H. Ramsey, P.D. Rostron, Measurement uncertainty from sampling and its role in validation of measurement procedures, Accred Qual. Assur. 29 (2024) 153, https://doi.org/10.1007/s00769-024-01575-0.
[80] J.D. Hernandez-Vásquez, E.C. de Oliveira, The importance of measurement uncertainty arising from the sampling process in conformity assessment: the case of fuel quality, Metrology 5 (2025), https://doi.org/10.3390/metrology5010007.
[81] E.C. de Oliveira, F.R. Lourenço, Risk of false conformity assessment applied to automotive fuel analysis: a multiparameter approach, Chemosphere 263 (2021), https://doi.org/10.1016/j.chemosphere.2020.128265.
[82] R.J.N.B. da Silva, F. Lourenço, D.B. Hibbert, Setting multivariate and correlated acceptance limits for assessing the conformity of items, Anal. Lett. 55 (2022), https://doi.org/10.1080/00032719.2022.2042549.
[83] T. Hartung, Artificial intelligence as the new frontier in chemical risk assessment, Front. Artificial Intell. 6 (2023), https://doi.org/10.3389/frai.2023.1269932.