Fit-for-purpose risks in conformity assessment of a substance or material – A case study of synthetic air

Fit-for-purpose risks in conformity assessment of a substance or material – A case study of synthetic air

Francesca R. Pennecchi[a], Ilya Kuselman[b,*], D. Brynn Hibbert[c], Michela Sega[a], Francesca Rolle[a] and Vladimir Altshul[d]

[a] Istituto Nazionale di Ricerca Metrologica (INRIM), Strada delle Cacce 91, 10135 Turin, Italy

[b] Independent Consultant on Metrology, 4/6 Yarehim St., 7176419 Modiin, Israel

[c] School of Chemistry, UNSW Sydney, Sydney NSW 2052, Australia

[d] Maxima Ltd, 10 Haogen St., 7714101 Ashdod, Israel

---

* Corresponding author. Tel.: +972-50-6240466

E-mail address: ilya.kuselman@bezeqint.net (I. Kuselman).

4/6 Yarehim St., Modiin, 7176419 Israel

ABSTRACT

A technique is described for evaluation of the fit-for-purpose risks in conformity assessment of the chemical composition of a substance or material, based on a multivariate Bayesian approach. The approach takes into account measurement uncertainty, correlation and the mass balance constraint. Two datasets related to synthetic air (provided as electronic supplementary material to this paper) were studied. The first dataset was from an industrial factory producing routinely medicinal synthetic air according to the European Pharmacopoeia. The second dataset was from the National Metrology Institutes which participated in key comparison CCQM-K120 "Carbon dioxide at background and urban level". The fitness for purpose of the preparation of synthetic air was interpreted as total risks of false decisions on the conformity of the air composition to the tolerance limits of the contents of its main components. Calculations of these risks were performed with code written in the R programming environment.

Keywords:

Conformity assessment; Synthetic air; Measurement uncertainty; Mass balance constraint; Risk of false decisions; Fitness for purpose

1. Introduction

Conformity assessment of the chemical composition of a substance or material provides evidence that specified requirements of concentrations or contents of the composition components are fulfilled [1]. A standard describing technical conditions of a multicomponent material includes specifications for its chemical composition, usually stated as limits of the tolerance interval of the actual ('true') content c_i of the i-th component, $i = 1, 2, \ldots, n$. Conformity assessment of an item (a material batch or lot) is based on comparing the measured content $c_{i\mathrm{m}}$ with the corresponding upper limit $T_\mathrm{U}(c_i)$ and lower limit $T_\mathrm{L}(c_i)$ of the tolerance (specification) interval T_i. Since any $c_{i\mathrm{m}}$ value has an associated standard measurement uncertainty $u(c_{i\mathrm{m}})$ [2, 3], the upper limit $A_\mathrm{U}(c_{i\mathrm{m}})$ and lower limit $A_\mathrm{L}(c_{i\mathrm{m}})$ of the acceptance interval A_i for measurement results may be used in place of the tolerance limits. In these cases, decisions (does the test item conform or not?) are based on comparing measured content values $c_{i\mathrm{m}}$ with the acceptance limits, which differ from the tolerance limits by a guard band (grey zone) proportional to $u(c_{i\mathrm{m}})$. When tolerance limits have been determined already taking into account measurement uncertainty, acceptance limits and tolerance limits coincide.

Measurement uncertainty causes risks of false (incorrect) decisions on the conformity of an item. The probability of accepting a batch, when it should have been rejected, is called the 'consumer's risk', whereas the probability of falsely rejecting a conforming batch is the 'producer's risk'. For a particular batch under test, they are referred to as the 'specific consumer's risk' and the 'specific producer's risk', respectively. The risks of an incorrect conformity assessment of a batch randomly drawn from a statistical population of such batches are the 'global consumer's risk' and the 'global producer's risk', respectively, as they characterize the batch production globally [4].

In general, a component-by-component evaluation of the risks of false decisions in the conformity assessment of a substance or material is not complete, as it does not give an answer to the question of the probability of a false decision on conformity of the product as a whole. When conformity assessment for each i-th component of a batch is successful (i.e., the *particular* specific or global risks are small enough), the total probability of a false decision concerning the batch as a whole (the multivariate *total* specific or *total* global risk) might still be significant. Evaluation

of the total risks is detailed in IUPAC/CITAC Guide [5] based on a multivariate Bayesian approach.

The term ‘content’ was used in this Guide for a quantity of a component of an item subject to conformity assessment (amount of substance, mass, volume, number of entities) expressed per unit mass of the item [5, Sec. 1.2]. In the present work the term ‘content’ and its symbol c are used as a generic term and symbol for both extensive and intensive measurands.

When the regulated components’ contents of a batch are subject to the mass balance constraint $\sum_i^n c_i$ = 100 or 1 or another constant value (e.g., for standard gas mixtures [6, 7]), they are intrinsically correlated. This so-called ‘spurious’ correlation is observed in addition to other possible natural and/or technological correlations between the components’ contents. All correlations may influence the understanding of test results (c_{im} with associated $u(c_{\mathrm{im}})$) and the evaluation of risks of false decisions in the conformity assessment. The circumstances mentioned above require appropriate modelling for the multivariate c_i distribution in different batches (prior distribution) and the multivariate c_{im} distribution in the same batch under test (likelihood function). Monte Carlo simulations in the R programming environment have been used for modelling in previous case studies [8-10].

When risks are evaluated, the question remains, what are the risk values that may satisfy both a producer of a substance or material and its consumer? An answer to such a question may be based on the concept of ‘fitness for purpose’ (fitness for intended use) [11]. Fitness for purpose describes acceptable quality of the fulfilment of specifications or stated outcomes [12]. It is widely used in accreditation [13], education [14, 15], engineering [16], validation of analytical chemical methods [17], proficiency testing [18], evaluation of measurement uncertainty arising from sampling [19, 20] and many other fields. If the risks are unacceptably great (i.e., not fit for purpose), they can be decreased by reducing measurement uncertainties of the measured values, as proposed for univariate global risks in JCGM 106 [4, Sec. 9.5.6]. In the Guide to decision-making and conformity assessment [21, Sec. 6.1] this idea was applied for conformance probability in a healthcare bivariate study of skin cream friction and adhesion.

Fit-for-purpose total risk values can be used for setting acceptable multivariate acceptance limits as proposed in ref. [22] and IUPAC/CITAC Guide [5, Sec. 5.6]. It is noted that setting comprehensive acceptance limits requires a study of not only producer’s and consumer’s risks as

probabilities depending on measurement uncertainty, but also economic, safety and/or other impacts of related false decisions [23], which are not discussed here.

The objective of the present paper is the development of a technique for evaluation of the fit-for-purpose risks in conformity assessment of chemical composition of a substance or material, based on a multivariate Bayesian approach, taking into account measurement uncertainty, correlation and the mass balance constraint.

As a case study two datasets related to synthetic air were analysed: 1) from the industrial factory Maxima [24], a routine producer of medicinal synthetic air according to the European Pharmacopoeia (EP) [25], and 2) from National Metrology Institutes (NMIs) that participated in the key comparison CCQM-K120 "Carbon dioxide at background and urban level" [26].

Synthetic air is a mixture of nitrogen and oxygen (and some other minor components or impurities) that is used as 'zero gas' in maintenance and calibration of test equipment for environmental monitoring, and as 'balance gas' in calibration mixtures. Since synthetic air contains oxygen, it is used for burning, respiration of plants and animals, decay and industrial oxidations (e.g., in metallurgical processes, pneumatic drills and plasma cutting), as well as in atomic absorption flame spectrometry and for flame ionization detectors in a laboratory [27, 28]. Other applications are for medical purposes [25, 29, 30].

2. Experimental

2.1. Medicinal synthetic air

Test results of N_{Max} = 316 lots of the EP medicinal synthetic air produced at Maxima in 2020 were used as a dataset for the case study of total risks. Each lot consisted of containers of the same volume (16 cylinders as a rule), prepared by evacuation and filled under specified pressure with pure dry oxygen and then nitrogen, using the same pipe-line in the same conditions, simultaneously. Both oxygen and nitrogen are produced at Maxima by air separation based on a cryogenic distillation process [31]. The filled cylinders are rolled to ensure homogeneity of the gas mixture in them. The factory tests one container from a lot for conformity assessment to the EP specifications before the product is transported to the consumer.

The dataset includes measured values of the components' contents in the N_{Max} lots, expressed as volume fractions of nitrogen ϕ_{1m} cL/L, oxygen ϕ_{2m} cL/L, and water vapour ϕ_{3m} µL/L, where cL is 10^{-2} L and µL is 10^{-6} L. It is provided in the electronic supplementary material to this paper (RawData_Maxima.txt file). Note that the units of volume fraction, % *V*/*V* and ppm *V*/*V*, used in the European Pharmacopeia, are replaced in this paper by cL/L and µL/L, respectively, to be consistent with IUPAC terminology and SI unit prefixes.

2.1.1. Specification and acceptance limits

Medicinal synthetic air is defined in the European Pharmacopoeia as a mixture of nitrogen (i =1) and oxygen (i = 2), where the volume fraction of oxygen is 95.0 % to 105.0 % of the nominal value which is between 21.0 cL/L to 22.5 cL/L. These specifications can be interpreted as the acceptance limits of measured oxygen volume fraction ϕ_{2m}, i.e., $A_L(\phi_{2m})$ = 21.0 cL/L and $A_U(\phi_{2m})$ = 22.5 cL/L, while the tolerance limits of actual ϕ_2 value are $T_L(\phi_2)$ = (95.0/100)·21.0 cL/L= 20.0 cL/L and $T_U(\phi_2)$ = (105.0/100)·22.5 cL/L= 23.6 cL/L.

Water vapour is specified in the European Pharmacopoeia as an impurity for which the maximum actual volume fraction ϕ_3 should be 67 µL/L. That means the acceptance limits of measured water vapour volume fraction ϕ_{3m} are $A_L(\phi_{3m}) = T_L(\phi_3)$ = 0 µL/L and $A_U(\phi_{3m}) = T_U(\phi_3)$ = 67 µL/L. Note that, even achieving $T_U(\phi_3)$, water volume fraction ϕ_3 is negligible in comparison with the volume fraction ϕ_2 of oxygen, being four orders of magnitude smaller.

Therefore, the mass balance for medicinal synthetic air $\sum_{i=1}^{3} \phi_i = 100$ cL/L can be simplified to $\phi_1 + \phi_2$ = 100 cL/L, from which nitrogen volume fraction is ϕ_1 = (100 – ϕ_2) cL/L. Hence, the acceptance limits of measured nitrogen volume fraction ϕ_{1m} are $A_U(\phi_{1m})$ = (100 – 21.0) cL/L = 79.0 cL/L and $A_L(\phi_{1m})$ = (100 – 22.5) cL/L = 77.5 cL/L. The tolerance limits of actual nitrogen volume fraction ϕ_1 are $T_L(\phi_1)$ = (100 – 23.6) cL/L = 76.4 cL/L and $T_U(\phi_1)$ = (100 – 20.0) cL/L = 80.0 cL/L.

Note also that water volume fraction ϕ_3, although insignificant for the mass balance, is still important for quality of the air as a product and must be taken into account in its conformity assessment.

2.1.2 Test methods and measurement uncertainties

Oxygen volume fractions are measured with a portable gas analyser Servomex MiniMP 5200 equipped with a high-performance sensor which is based on the paramagnetic susceptibility of the oxygen molecule. This physical property distinguishes oxygen from most other gases [32]. Accuracy of results provided by the instrument is ± 0.02 cL/L; resolution 0.01 cL/L; output fluctuation ± 0.01 cL/L; zero drift per week (calibration interval) less than 0.15 cL/L. The drift is the dominant component of the measurement uncertainty here. Assuming drift values have a rectangular distribution, the standard measurement uncertainty is $u(\phi_{2m})$ = 0.15/√3 cL/L= 0.09 cL/L.

Water vapour volume fractions are measured with a Shaw Dew Point SADP-Red Spot Meter providing direct indication in dew point temperature [33], which can be converted into water vapour volume fraction, µL/L [34]. Accuracy of results provided by the instrument is ± 3 °C, equal to about ± 1 µL/L; resolution is 1 °C (about 0.3 µL/L); output fluctuation is less than 1 °C; a drift is negligible as the analyser is equipped with an automatic calibration control. Assuming resolution and output fluctuation as negligible, and a rectangular distribution of the accuracy parameter, the standard measurement uncertainty can be evaluated as $u(\phi_{3m})$ = 1/√3 µL/L= 0.6 µL/L.

Standard measurement uncertainty of nitrogen volume fraction calculated as $\phi_{1m} = (100 - \phi_{2m})$ cL/L is the same as of oxygen volume fraction, i.e., $u(\phi_{1m}) = u(\phi_{2m})$ = 0.09 cL/L.

Note that the terminology of the manuals of the measuring instruments [32] and [33] differs from that in the International Vocabulary of Metrology (VIM) [35]. In particular, ‘accuracy’ as used in the manuals is interpreted as ‘trueness’ in the VIM.

2.2. Synthetic air for CCQM-K120

The CCQM-K120 key comparison was designed to evaluate the level of compatibility of NMIs’ preparative capabilities for carbon dioxide in air. Synthetic air was used as the balance gas (dry air matrix) for preparation of the measurement standards of carbon dioxide at background and urban level [26] with the carbon dioxide amount fraction (380 – 480) µmol/mol in CCQM-K120.a and (480 – 800) µmol/mol in CCQM-K120.b, respectively. Also scrubbed ‘real’ air was used by some NMIs as the balance gas, not discussed here. Fifteen NMIs were involved in this comparison,

twelve of them prepared 23 gas measurement standards (cylinders) with the synthetic air containing carbon dioxide (380 – 480) µmol/mol. Note that a standard containing carbon dioxide (480 ± 10) µmol/mol was applicable for both parts of CCQM-K120, when requirements of the synthetic air were satisfied. For example, the standard prepared by INRIM (the Italian NMI) for CCQM-K120.b also satisfied the requirements for CCQM-K120.a. Thus, in the present work we refer to the dataset of the synthetic air properties of the $N_{CCQM} = 23$ cylinders and the requirements of CCQM-K120.a for this air.

The air was synthesized at the NMIs gravimetrically, blending purchased pure gases, according to ISO 6142-1 [36]. The dataset including acronyms of the twelve NMIs and their measured (assigned) values of components' contents in the N_{CCQM} cylinders, expressed in amount fractions (mol/mol) of nitrogen x_{1m}, oxygen x_{2m} and argon x_{3m}, is provided as electronic supplementary material to this paper (RawData_CCQM.txt file).

2.2.1. Specification and acceptance limits

Nitrogen, oxygen and argon were defined in the protocol of the CCQM-K120.a comparison [37] as the main components of the air with the following tolerance limits of the amount fractions, mol/mol:

i =1) nitrogen $T_L(x_1) = 0.7804 \le x_1 \le 0.7814 = T_U(x_1)$;

i =2) oxygen $T_L(x_2) = 0.2088 \le x_2 \le 0.2098 = T_U(x_2)$;

i =3) argon $T_L(x_3) = 0.0089 \le x_3 \le 0.0097 = T_U(x_3)$.

Since the tolerance limits were set based on the published measured values of the component amount fractions in ambient (real) air, the acceptance limits $A_L(x_{im})$ and $A_U(x_{im})$ are considered here equal to the respective tolerance limits $T_L(x_i)$ and $T_U(x_i)$.

In addition to carbon dioxide (i = 4) 380 ± 10 µmol/mol in one cylinder of a comparison participant, and 480 ± 10 µmol/mol in a second cylinder, the air could contain nitrous oxide (i = 5) and methane (i = 6) as greenhouse gases, with tolerance limits from zero to 1900 nmol/mol and 330 nmol/mol, respectively. As µmol is 10^{-6} mol and nmol is 10^{-9} mol, amount fractions of these components do not influence the mass balance $\sum_{i=1}^{6} x_i = 1$ mol/mol. For simplicity they were not taken into account further.

2.2.2. Methods of the gas standard preparation and measurement uncertainties

Details of preparation of the gas standards and evaluation of measurement uncertainties are described in the measurement reports of the corresponding twelve NMIs [37, Annex 5]. For example, at INRIM the gas mixture was prepared in 5 L aluminium cylinders, which were preconditioned by evacuation and heating. The preparation was performed by adding sequentially the calculated masses of the air components, in the following order: carbon dioxide from a diluted parent mixture with nitrogen (5005.4 µmol/mol of CO_2), a parent mixture of argon with oxygen (0.042703 mol/mol of Ar), and pure nitrogen. Each addition was followed by precision weighing of the cylinder using a mass comparator. The weighing double-substitution scheme A-B-B-A was applied, where A was the sample cylinder, and B a reference cylinder, kept empty during the entire procedure. Calibrated masses added to the lighter cylinder, keeping the mass difference between the two cylinders within 1 g for optimizing the mass comparator performance and obtaining smaller gravimetric uncertainties. After preparation the cylinders were rolled to homogenize the gases.

The main contributions to the uncertainty of the amount fractions of the synthetic air components from the gravimetric preparation were of the weighted masses of the parent mixtures, including buoyancy correction, the molar masses of the purchased gases and their purity. Some NMIs, e.g., NPL (UK) and NPLI (India) also measured the obtained component amount fractions using physical-chemical methods (e.g., gas chromatography with flame ionization detection), combining the uncertainty components from the gravimetric preparation and the measurement, thus reporting increased uncertainty values. Therefore, the standard measurement uncertainties (equal to ½ of the reported expanded uncertainties which used a coverage factor of 2 [37], listed in the RawData_CCQM.txt file) for nitrogen $u(x_{1m})$ varied widely in the interval from 0.000002 mol/mol at KRISS (Korea) to 0.001105 mol/mol at NPLI. A similar interval of standard measurement uncertainty values $u(x_{2m})$ for oxygen was (0.000002 – 0.000295) mol/mol, and the interval of $u(x_{3m})$ values for argon was (0.000001 – 0.000015) mol/mol.

Note that the standard uncertainty values for nitrogen and oxygen reported by NPLI, $u(x_{1m})$ = 0.001105 mol/mol and $u(x_{2m})$ = 0.000295 mol/mol, are significantly greater than the respective values from the other NMIs.

3. Modelling and calculation

3.1. Background

A Bayesian model of the multivariate pdf of n components' contents in a synthetic air is expressed by the following equation:

$$g(\boldsymbol{c} \mid \boldsymbol{c}_\mathrm{m}) = C g_0(\boldsymbol{c}) h(\boldsymbol{c}_\mathrm{m} \mid \boldsymbol{c}), \tag{1}$$

where $\boldsymbol{c} = [c_1, c_2, \ldots, c_n]$ and $\boldsymbol{c}_\mathrm{m} = [c_{1\mathrm{m}}, c_{2\mathrm{m}}, \ldots, c_{n\mathrm{m}}]$ are vectors of the actual ('true') values c_i and measured values $c_{i\mathrm{m}}$, respectively, for $i = 1, 2, \ldots, n$; $g(\boldsymbol{c} \mid \boldsymbol{c}_\mathrm{m})$ is the posterior pdf representing post-measurement knowledge about the distribution of actual values c_i; C is a normalizing constant; $g_0(\boldsymbol{c})$ is the prior pdf, expressing pre-measurement knowledge about the distribution of c_i, taking into account correlations between c_i; and $h(\boldsymbol{c}_\mathrm{m} \mid \boldsymbol{c})$ is the likelihood function expressing knowledge about the distribution of measured values $c_{i\mathrm{m}}$ (vector $\boldsymbol{c}_\mathrm{m}$), involving the measurement uncertainties and correlations between $c_{i\mathrm{m}}$, for a given actual values c_i (vector $\boldsymbol{c}$) [5]. Note again that the term 'content' and its symbol c are generic here for either volume fraction ϕ or amount fraction x.

The prior pdf is modelled using a theoretical pdf that is fitted to a dataset of actual values c_i characterizing the production process, i.e., the measured values of the chemical composition of produced synthetic air, accumulated during a specified time of production. Adequacy of the prior as the theoretical distribution of c_i is proved by testing its goodness-of-fit. When more than one theoretical distribution is adequate, the simplest is preferable as the prior.

The likelihood function is recovered from knowledge about the measurement uncertainty (pdf of the measured values $c_{i\mathrm{m}}$ at the same actual values c_i in the same air sample) available from the document on the measurement process and the measurement procedure.

The posterior pdf is derived as the normalized product of prior and likelihood. It contains an updated state of knowledge about the product when further measured values $c_{i\mathrm{m}}$ are obtained (after accumulation of the dataset for modelling the prior pdf) and predicts corresponding further actual content values c_i. Note that a posterior pdf modelled in this way is applicable as long as the

production and measurement conditions (reflected in the prior and likelihood, respectively) are not changed.

The total global consumer's risk R_c and the total global producer's risk R_p are, respectively [8]:

$$R_\mathrm{c} = \int_{T^\mathrm{c}} \int_A \; g_0(\boldsymbol{c}) h\big(\boldsymbol{c}_\mathrm{m} \,|\, \boldsymbol{c}\big)\, d\boldsymbol{c}_\mathrm{m} d\boldsymbol{c} \quad \text{and} \quad R_\mathrm{p} = \int_T \; \int_{A^\mathrm{c}} g_0(\boldsymbol{c}) h\big(\boldsymbol{c}_\mathrm{m} \,|\, \boldsymbol{c}\big)\, d\boldsymbol{c}_\mathrm{m} d\boldsymbol{c}, \tag{2}$$

where T is the tolerance/specification domain $T_1 \times T_2 \times \ldots \times T_n$, A is the acceptance domain $A_1 \times A_2 \times \ldots \times A_n$, and 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 $c_{i\mathrm{m}}$ is performed within A. The subscript "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 consumer's risk R_c^* and the total specific producer's risk R_p^* are, respectively [9]:

$$R_\mathrm{c}^* = 1 - \int_T \; g\big(\boldsymbol{c} \,|\, \boldsymbol{c}_\mathrm{m}\big) d\boldsymbol{c} \quad \text{when } \boldsymbol{c}_\mathrm{m} \text{ is in } A, \text{ and}$$

$$R_\mathrm{p}^* = \int_{T_1} \ldots \int_{T_\nu} \int_0^{100} \ldots \int_0^{100} g\big(\boldsymbol{c} \,|\, \boldsymbol{c}_\mathrm{m}\big) \mathrm{d}\boldsymbol{c} \quad \text{when } c_{i\mathrm{m}},\ 1 \le i \le \nu, \text{ are outside } A. \tag{3}$$

Here, R_c^* is the probability that at least one of corresponding actual content values c_i of synthetic air components is actually outside its tolerance interval T_i, when all the measured content values $c_{i\mathrm{m}}$ are within their acceptance intervals A_i (false conforming). Thus, R_c^* is equal to one minus the probability that all c_i are within the tolerance domain T when all $c_{i\mathrm{m}}$ conform, i.e., are within the acceptance domain A.

Symbol ν in Eq. (3) for R_p^* indicates the number of those air components, $1 \le \nu \le n$, whose measured content values $c_{i\mathrm{m}}$ are outside their acceptance intervals A_i. Hence, the vector $\boldsymbol{c}_\mathrm{m}$, being out of the acceptance domain A for those ν components of the tested product, is rejected as non-conforming. For simplicity, and without losing generality, the measured values $c_{i\mathrm{m}}$ outside their acceptance intervals are the first ν values. Given that these ν measured values do not conform, R_p^*

is the probability that the corresponding actual values are actually all inside their tolerance intervals, hence the product does satisfy its specifications and its rejection is a false decision.

When contents of one only particular component are under control/testing, Eqs (1)-(3) reduce to the univariate counterparts described in JCGM 106 [4]. If the contents of the components are independent one from each other, hence Eqs (2)-(3) simply involve the evaluation of particular risks for each component [5].

3.2. Analysis of the datasets

3.2.1. Medicinal synthetic air

Empirical distributions of the test results are characterized in Table 1, where $\phi_{i\min}$ and $\phi_{i\max}$ are the minimum and the maximum measured volume factions of the i-th component in the dataset, respectively; $m(\phi_{im})$ and $s(\phi_{im})$ are the sample mean and the standard deviation, respectively.

Table 1

Table 1. Parameters of the empirical distributions of the measured component volume fractions in the medicinal synthetic air and of the fitted theoretical contaminated normal distributions (Eq. 4).

Parameters of the empirical distributions*						Parameters of the fitted theoretical distributions†							
i	Comp.	$\phi_{i\min}$	$\phi_{i\max}$	$m(\phi_{im})$	$s(\phi_{im})$	μ_{i1}	σ_{i1}	w_{i1}	μ_{i2}	σ_{i2}	w_{i2}	D_i	P_i
1	Nitrogen	77.5	79.0	78.4	0.4	78.4	0.4	0.1	78.9	0.04	0.9	0.063	0.16
2	Oxygen	21.0	22.5	21.6	0.4	21.1	0.04	0.9	21.6	0.4	0.1	0.063	0.16
3	Water	0.2	2.5	1.2	0.6	0.6	0.2	0.6	1.5	0.4	0.4	0.064	0.15

* The unit for volume fractions of nitrogen and oxygen (i = 1 and 2) is cL/L, while for water vapour (i = 3) it is µL/L.

† μ and σ are the means and standard deviations of the theoretical distributions having the same units as the respective parameters of the empirical distributions; the weights w_{i1} and w_{i2}, and parameters of the Kolmogorov-Smirnov test of their goodness-of-fit, D_i and P_i, are dimensionless.

The standard deviations $s(\phi_{1m})$ and $s(\phi_{2m})$ of measured nitrogen and oxygen volume fractions are almost five times greater than the corresponding standard measurement uncertainty $u(\phi_{1m}) = u(\phi_{2m})$. This difference in magnitude is observed because the standard deviations are influenced by the variability of conditions of the technological process during the year of the air production when the dataset was accumulated, as well as by measurement uncertainty. However, water vapour is an impurity and the standard deviation $s(\phi_{3m})$ is approximately the same as the measurement uncertainty $u(\phi_{3m})$.

The Pearson correlation coefficient between measured values of oxygen and nitrogen is $r_{12} = -1$ by definition, since $\phi_{1m} = (100 - \phi_{2m})$ cL/L. The correlation coefficient between measured values of oxygen and water vapour, calculated from the dataset, is $|r_{23}| = 0.049$. It is statistically insignificant, as the critical value for such coefficient at the 0.95 level of confidence for a two-sided *t*-test for a sample size of 316 pairs of measured values, is 0.110 [38].

As the raw gases for preparation of synthetic air, pure nitrogen and oxygen, are produced at Maxima from ambient air (having an oxygen volume fraction not achieving 21.0 cL/L [26]), there is a need to set the average oxygen volume fraction in the synthetic air closer to the lower acceptance limit. Therefore, the default partial pressure of oxygen in the cylinders was calculated to maintain an average oxygen volume fraction value of about 21.6 cL/L. When the measured oxygen volume fraction ϕ_{2m} in the lot of cylinders violated the lower acceptance limit $A_L(\phi_{2m}) = 21.0$ % cL/L, the oxygen partial pressure was increased to get an amount somewhat greater than 21.0 cL/L. Thus, the empirical ϕ_{2m} distribution in the acceptance interval (21.0 – 22.5) cL/L, shown in Fig. 1a as a histogram, is skewed to the right and has two maximums: 1) at an oxygen volume fraction of about 21.6 cL/L as the average corresponding to the default oxygen pressure, and 2) at about 21.1 cL/L, i.e., close to the lower acceptance limit.

Fig. 1

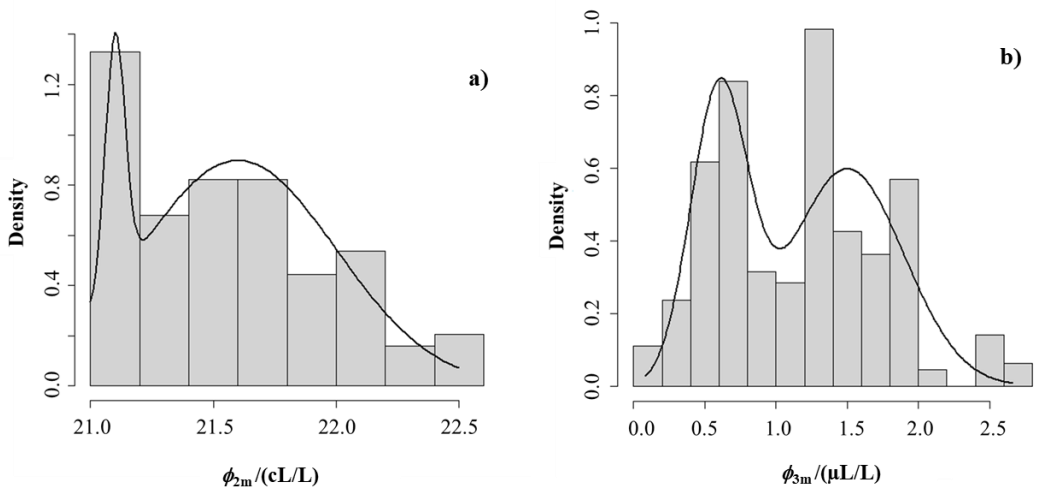

Figure 1. Histograms of measured values (a) ϕ_{2m} of oxygen volume fraction, and (b) ϕ_{3m} of water vapour volume fraction, in medicinal synthetic air. The solid lines indicate the theoretical contaminated normal pdfs. The ordinate is related to the probability density of the pdfs (relative frequency of the data expressed in 1/unit of the abscissa).

Summarizing the data, the histogram in Fig. 1a was interpreted as a contaminated normal distribution [39, 40], i.e., a mixture of two normal distributions, one with the maximum of the probability density function (pdf) at 21.6 cL/L of oxygen, and the second one at 21.1 cL/L. A specular histogram corresponds to nitrogen measured fractions $\phi_{1m} = (100 - \phi_{2m})$ cL/L.

The available measured values of water vapour volume fraction were significantly discretized because of the resolution of the water analyser of 0.3 µL/L (Sec. 2.1.2). In order to find a continuous pdf to model the data, a uniform noise was added to the observed data. In this way, 316 new water values were generated, each one drawn from a uniform pdf centred on the original value ϕ_{3m} and randomly deviated from it within the range of the analyser resolution of 0.3 µL/L. The sample mean and standard deviation of the new data were practically the same as those of the original data in Table 1.

The histogram of the generated data is shown in Fig. 1b. A contaminated normal distribution was used again for modelling the data, combining the following two distributions: 1) one having a pdf maximum at 0.6 µL/L, that can be related to the lots of cylinders when the default pressures were used, and 2) a distribution with its maximum at 1.5 µL/L that is related to cylinders in which the initial partial oxygen pressure violated the lower acceptance limit $A_L(\phi_{2m}) = 21.0$ cL/L

requiring the addition of oxygen. Note that the second maximum is caused as more operations lead to more possibilities for water vapour to penetrate into cylinders with the air.

Hence, for both oxygen and water vapour volume fractions, the theoretical production distribution is assumed to be a contaminated normal distribution $f(\phi_i)$, which is a combination of two normal distributions $f_1(\phi_i)$ and $f_2(\phi_i)$ with non-negative weights $w_{i1} = (1 - w_{i2})$ and w_{i2}, respectively, such that the sum of the weights is equal to 1:

$$f(\phi_i) = w_{i1}\, f_1(\phi_i;\, \mu_{i1},\, \sigma_{i1}) + w_{i2}\, f_2(\phi_i;\, \mu_{i2},\, \sigma_{i2}), \tag{4}$$

where the weights and other pdf parameters for the present case study are shown in Table 1.

Note that the pdf of the nitrogen volume fraction, being a mirror of the pdf of the oxygen volume fraction, refers to nitrogen volume fractions that are a complement to 100 cL/L of the corresponding oxygen volume fractions, according to the mass balance constraint. For example, when oxygen volume fractions $\phi_2 = w_{21}\mu_{21} + w_{22}\mu_{22} = 0.1{\cdot}21.1$ cL/L + $0.9{\cdot}21.6$ cL/L = 21.6 cL/L, the corresponding volume fraction of nitrogen is $\phi_1 = w_{11}\mu_{11} + w_{12}\mu_{12} = 0.9{\cdot}78.4$ cL/L + $0.1{\cdot}78.9$ cL/L = 78.4 cL/L, and their sum is 100 cL/L.

R functions for calculations on contaminated normal pdfs are available, for example, in ref. [41]. Results of one sample Kolmogorov-Smirnov test of goodness-of-fit of the theoretical production contaminated normal distribution and the respective empirical one (from the dataset) with alternative two-sided hypothesis [38, 42] are presented in Table 1. They are very similar, since the oxygen addition and related technical reasons leading to those contaminated normal distributions for the three components were the same. Based on these results, the hypotheses of goodness-of-fit were not rejected at the 0.95 level of confidence as for oxygen and nitrogen, as for water vapour. The theoretical contaminated normal pdfs are shown in Fig. 1 with solid lines.

3.2.2. Synthetic air for CCQM-K120

Distributions of the accumulated measured values are characterized in Table 2, where $x_{i\min}$ and $x_{i\max}$ are the minimum and the maximum of the i-th component amount fraction in the dataset, respectively; $m(x_{i\mathrm{m}})$ and $s(x_{i\mathrm{m}})$ are the interlaboratory mean and standard deviation, respectively.

Table 2

Calculated Pearson correlation coefficients r_{ij} (i, j = 1, 2, 3, $i \neq j$) between measured components' amount fractions are presented in Table 3. The critical two-sided value of the correlation coefficients for (23 – 2) = 21 degrees of freedom and 0.95 level of confidence is 0.413 [38].

Table 3

As $|r_{i3}| \leq 0.413$ (i = 1, 2) in Table 3, the correlation of the argon amount fractions with the nitrogen and oxygen amount fractions is considered insignificant. However, correlation of the nitrogen and oxygen amount fractions (having a 'spurious' origin caused by the mass balance constraint [8]) is statistically significant, since $|r_{12}| = 0.767 > 0.413$.

The empirical distributions of the measured component amount fractions x_{1m}, x_{2m} and x_{3m} are shown in Fig. 2a, 2b and 2c, respectively, as histograms.

Fig. 2

Table 2. Parameters of the distributions of the measured component amount fractions in synthetic air for CCQM-K120 and their fitting by theoretical normal distributions.

i	Comp.	$x_{i\min}$/ (mol/mol)	$x_{i\max}$/ (mol/mol)	$m(x_{im})$/ (mol/mol)	$s(x_{im})$/ (mol/mol)	D_i	P_i
1	Nitrogen	0.7800	0.7820	0.7809	0.00046	0.129	0.84
2	Oxygen	0.2088	0.2102	0.2094	0.00036	0.145	0.72
3	Argon	0.0089	0.0096	0.0093	0.00015	0.126	0.86

Table 3. Matrix of Pearson correlation coefficients r_{ij} of the component measured amount fractions in synthetic air prepared in CCQM-K120.

i / *j*	Comp.	Nitrogen	Oxygen	Argon
		1	2	3
1	Nitrogen	1.000	-0.767	-0.348
2	Oxygen	-0.767	1.000	-0.162
3	Argon	-0.348	-0.162	1.000

Results of the one sample Kolmogorov-Smirnov test of goodness-of-fit of a theoretical normal distribution and the empirical distribution (the dataset) with the alternative two-sided hypothesis are presented in Table 2. Based on these results, the hypotheses of goodness-of-fit were not rejected at the 0.95 level of confidence. The fitted theoretical normal pdfs are shown in Fig. 2 as solid lines.

Note that the distributions of the components' amount fractions shown in Fig. 2 are considered as the marginal distributions of a multivariate pdf. Such a multivariate pdf is used in the following sections for modelling the prior pdf of the compositions of synthetic air in CCQM-K120. Its covariance matrix is in Table 4, where variances $s^2(x_{im})$ (values of $s(x_{im})$ are in Table 2) are the diagonal elements, and covariances $cov_{ij} = r_{ij}\, s(x_{im})\, s(x_{jm})$, $i \neq j$ (values of r_{ij} are in Table 3) are the off-diagonal elements.

Table 4

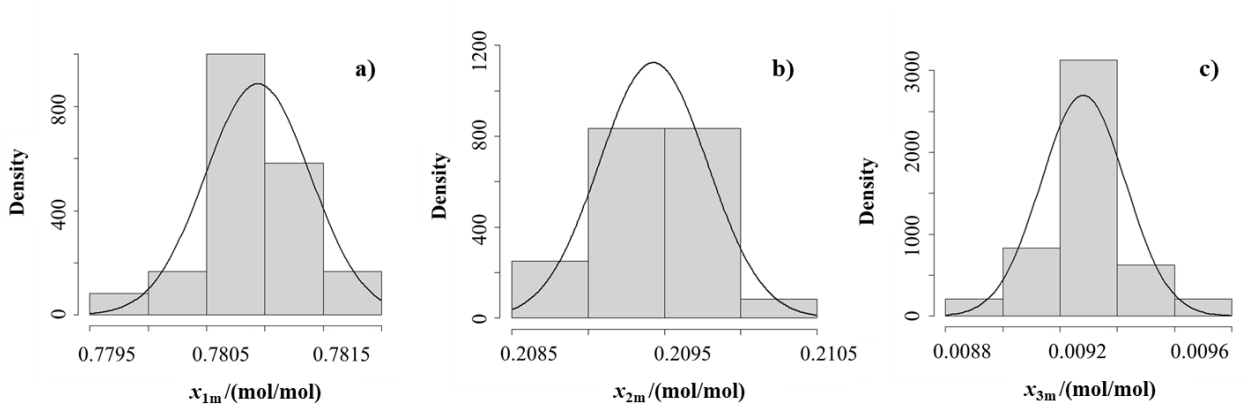

Figure 2. Histograms of measured amount fractions x_{1m} of nitrogen (a), x_{2m} of oxygen (b), and x_{3m} of argon (c) in the synthetic air for CCQM-K120. The solid lines indicate the fitted theoretical normal pdfs. The ordinate scale is described in the caption to Fig. 1.

Table 4. Covariance cov_{ij} matrix of the synthetic air data of CCQM-K120.

i / *j*	Comp.	Nitrogen	Oxygen	Argon
		1	2	3
1	Nitrogen	$2.12 \cdot 10^{-7}$	$-1.28 \cdot 10^{-7}$	$-2.42 \cdot 10^{-8}$
2	Oxygen	$-1.28 \cdot 10^{-7}$	$1.32 \cdot 10^{-7}$	$-8.88 \cdot 10^{-9}$
3	Argon	$-2.42 \cdot 10^{-8}$	$-8.88 \cdot 10^{-9}$	$2.29 \cdot 10^{-8}$

3.3. Modelling the risks

3.3.1. Prior pdfs and conformance probabilities

The univariate prior pdfs for the components of the medicinal synthetic air were modelled by the contaminated normal pdfs discussed in Sec. 3.2.1. The probability of conformance of the oxygen pdf, calculated as the fraction of $M = 1 \cdot 10^7$ Monte Carlo (MC) simulations of the events when the oxygen volume fractions ϕ_2 are within the tolerance interval $T_2(\phi_2)$, was equal to P_{conf} = 0.99997. As expected, this is also equal to the probability of conformance of the prior pdf modelling the nitrogen volume fractions ϕ_1. The probability of conformance for water volume fractions ϕ_3 is equal to one, since they are extremely far from their upper tolerance limit.

As the nitrogen volume fractions are calculated from the oxygen ones based on the mass balance, hence any bivariate pdf modelling for the two components actually degenerates into a univariate one, as in the case study of ref. [9] and Sec 3.1. The water vapour volume fractions are not correlated with the oxygen (and nitrogen) volume fractions and considered as independent from those. For these two reasons, the total probability of conformance for the medicinal synthetic

air as a multicomponent material is equal to the product of that for oxygen and for water vapour, hence coinciding with the probability of conformance of the oxygen (or nitrogen) volume fractions.

The multivariate prior pdf of the synthetic air for CCQM-K120 was modelled on the basis of the best available knowledge, described in Sec. 3.2.2. A total $M = 1 \cdot 10^7$ MC simulations of the actual air compositions $\boldsymbol{x} = [x_1, x_2, x_3]$ were performed as in the paper [8]. The x_i values were drawn from a multivariate normal distribution (truncated on the domain $[0, 1]^3$) with the location parameter equal to the mean vector $\boldsymbol{\mu} = [m(x_{1m}), m(x_{2m}), m(x_{3m})]$, and the scale parameter equal to the covariance matrix in Table 4. The values $m(x_{im})$ are available in Table 2.

Actual ('true') values have no uncertainties by definition [35] and the sum of actual amount fractions of nitrogen, oxygen and argon prepared in CCQM-K120 must be exactly equal to 1. Therefore, the data drawn from the multivariate truncated normal pdf were subjected to the closure operation:

$$clo\,(\boldsymbol{x}) = \left[\frac{x_1}{\sum_{i=1}^{3} x_i},\ \frac{x_2}{\sum_{i=1}^{3} x_i},\ \frac{x_3}{\sum_{i=1}^{3} x_i}\right]. \qquad (5)$$

Table 5

The resulting correlation matrix is given in Table 5.

Table 5. Matrix of correlation coefficients r_{ij} of the prior pdf of the component amount fractions in synthetic air prepared in CCQM-K120 (after the closure operation).

i / *j*	Comp.	Nitrogen	Oxygen	Argon
		1	2	3
1	Nitrogen	1.000	-0.919	-0.284
2	Oxygen	-0.919	1.000	-0.118
3	Argon	-0.284	-0.118	1.000

Comparing correlation matrices in Table 3 and Table 5, one can see that the correlation coefficient r_{12} related to amount fractions of nitrogen and oxygen was increased after the closure operation. The absolute values of the statistically insignificant correlation coefficients r_{13} and r_{23} between amount fractions of argon and other gases were even lower. The corresponding covariance matrix is shown in Table 6, where variance cov_{11} of nitrogen amount fractions, covariance cov_{13} between nitrogen and argon amount fractions, as well as covariance cov_{23} between oxygen and argon amount fractions decreased (as absolute values) with respect to their counterparts in Table 4.

Table 6

Table 6. Covariance cov_{ij} matrix of prior pdf for the synthetic air data of CCQM-K120 (after the closure operation).

i / *j*	Comp.	Nitrogen	Oxygen	Argon
		1	2	3
1	Nitrogen	$1.46 \cdot 10^{-7}$	$-1.29 \cdot 10^{-7}$	$-1.65 \cdot 10^{-8}$
2	Oxygen	$-1.29 \cdot 10^{-7}$	$1.36 \cdot 10^{-7}$	$-6.63 \cdot 10^{-9}$
3	Argon	$-1.65 \cdot 10^{-8}$	$-6.63 \cdot 10^{-9}$	$2.31 \cdot 10^{-8}$

The probability of conformance of the multivariate prior pdf before the closure operation, calculated as the fraction of M of the events when the simulated synthetic air compositions $\boldsymbol{x} = [x_1, x_2, x_3]$ are within the tolerance domain T, was $P_{conf} = 0.635$. After the closure operation, the probability decreased to $P_{conf} = 0.475$. These low values of conformance probability are caused by nitrogen contents out of the tolerance interval T_1 in five cylinders (OMH54, FB03747, JJ108862, PSM298266, PSM266468) and oxygen contents out of the tolerance interval T_2 in four cylinders (1029047, FB03744, PSM298266, PSM266468), in the dataset of $N_{CCQM} = 23$ cylinders, RawData_CCQM.txt file. Note that the correlation between nitrogen and oxygen amount fractions increased (in absolute value) after the closure operation, hence complicating the multivariate prior pdf and decreasing the conformance probability of the model based on the multivariate truncated normal pdf.

3.3.2. Likelihood function

The univariate oxygen and water vapour likelihood functions for the medicinal synthetic air were modelled using, respectively, a normal distribution and a normal distribution truncated at zero (in order to avoid simulation of negative water vapour values). The location parameters were equal to actual values ϕ_2 and ϕ_3 drawn from the corresponding prior pdfs, and scale parameters were equal to $u(\phi_{2m})$ and $u(\phi_{3m})$, respectively, defined in Sec. 2.1.2.

Modelling the multivariate likelihood function for a vector of amount fraction values $\boldsymbol{x}_\mathbf{m} = [x_{1m}, x_{2m}, x_{3m}]$ of the synthetic air for CCQM-K120, was based on $\boldsymbol{x}_\mathbf{m}$ recovering as $\boldsymbol{x}_\mathbf{m} = \boldsymbol{x} + \boldsymbol{e}_\mathbf{m}$, where $\boldsymbol{e}_\mathbf{m} = [e_{1m}, e_{2m}, e_{3m}]$ is the vector of measurement errors. The vector $\boldsymbol{e}_\mathbf{m}$ was modelled assigning a multivariate truncated normal pdf with zero expectation, while the vector of actual amount fraction values $\boldsymbol{x} = [x_1, x_2, x_3]$ was generated from the multivariate prior pdf [8].

The covariance matrix, used as the scale parameter of the truncated pdf associated with the vector $\boldsymbol{e}_\mathbf{m}$, contains as the diagonal elements the squared standard measurement uncertainties $u^2(x_{im})$ discussed in Sec. 2.2.2 and Sec. 3.2.2. The off-diagonal elements of this matrix are the covariance terms equal to products $cov_{ijlf} = r_{ij}\, u(x_{im})\, u(x_{jm})$ whose correlation coefficients r_{ij} are in Table 3. Subscript 'lf' in cov_{ijlf} means 'likelihood function'.

Since the likelihood function characterizes the measurement process with corresponding measurement uncertainties, the sum of the measured values in the vector $\boldsymbol{x}_\mathbf{m} = [x_{1m}, x_{2m}, x_{3m}]$ should not be equal to 1 exactly, and therefore, the closure operation is not applied here.

3.3.3. Posterior multivariate pdf

Once the prior pdf and likelihood function are modelled, the posterior pdf of the actual content values $\boldsymbol{c}$ in the synthetic air at the measured values of the components' contents $\boldsymbol{c}_\mathrm{m}$ can be calculated by Eq. (1) as the normalization of the product $g_0(\boldsymbol{c})h(\boldsymbol{c}_\mathrm{m} \mid \boldsymbol{c})$, where for the medicinal synthetic air and synthetic air for CCQM K120, c is the actual volume fraction value ϕ or the actual amount fraction x, respectively, and c_m is the measured volume fraction value ϕ_m or the measured amount fraction x_m, respectively.

Since the posterior pdf predicts further actual content values (after accumulation of the dataset 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 [10]. In other words, the sum of the calculated (predicted) actual component content values may differ from 100 % because any predicted value has its associated prediction uncertainty [43].

The code for calculations of the total global and the total specific risks, written in the R programming environment, is provided as electronic supplementary material to this paper (Synthetic_air_Rcode.txt file).

4. Results and discussion

4.1. Risks in conformity assessment of medicinal synthetic air

Any total risk for oxygen and nitrogen volume fractions in medicinal synthetic air, similar to the probability of conformance, simplifies to a particular risk relevant to either one of the two, as discussed in Sec. 3.1 and Sec. 3.3.1. Since the volume fractions of oxygen (or nitrogen) and water vapour are not correlated, the total risks relevant to oxygen and water vapour are equal to a combination of their particular risks [5]. The probability of conformance for water vapour volume fractions is equal to one (Sec 3.3.1), therefore a total risk for the medicinal synthetic air, equal to the product of the particular risks for oxygen and for water vapour, coincides with the particular risk for oxygen.

The particular global consumer's risk R_{c} related to the volume fractions of oxygen (and, therefore, of nitrogen) is equal to zero, while the particular global producer's risk is R_{p}= 0.0926. Both particular global consumer's and producer's risks related to the water volume fraction are zero. Therefore, the total global consumer's risk related to oxygen (or nitrogen) and water vapour is zero, according to the IUPAC/CITAC Guide [5], whereas the total global producer's risk, coinciding with the particular risk related to oxygen, is R_{p}= 0.0926.

The particular specific consumer's risks R_{c}^{*} related to oxygen and water vapour volume fractions, for $\phi_{2\mathrm{m}}$ and $\phi_{3\mathrm{m}}$ values within the corresponding acceptance intervals, are zero. Hence, also, the total specific consumer's risk, related to oxygen and water vapour, is zero [5].

The specific producer's risks R_{p}^{*} as a function of the measured oxygen volume fractions on the intervals from the lower acceptance limit to the corresponding tolerance limit (a) and similarly for the upper limits (b) are shown in Fig. 3.

Fig. 3

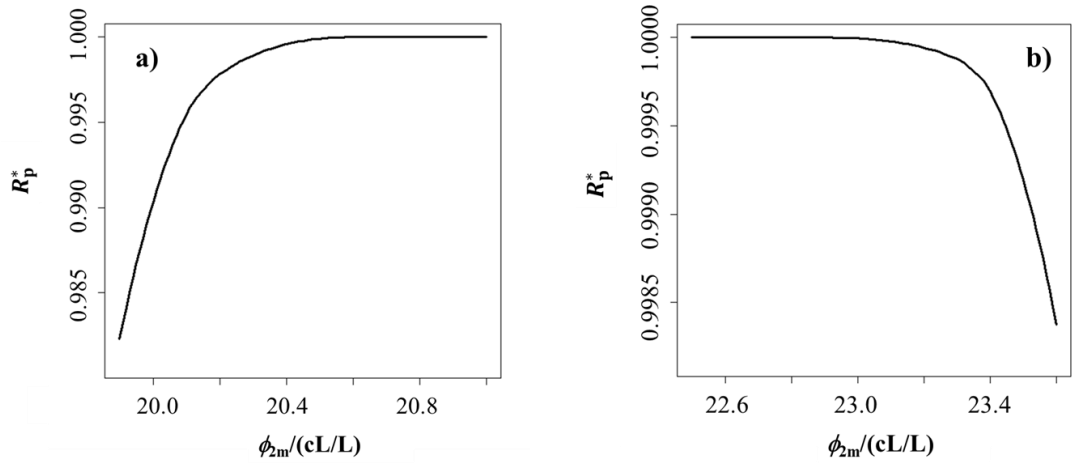

Figure 3. Dependence of the specific producer's risk R_p^* on the measured oxygen volume fractions ϕ_{2m} in medicinal synthetic air. The interval (20.0 - 21.0) cL/L, from the lower tolerance to the lower acceptance limits, is on plot (a), while the interval (22.5 - 23.6) cL/L, from the upper acceptance to the upper tolerance limits is on plot (b).

It is clear from Fig. 3 that the acceptance limits set in the European Pharmacopoeia are extremely reliable: the producer's risk decreases only when the oxygen volume fraction is less than the lower acceptance limit about 0.4 cL/L and greater than the upper acceptance limit about 0.4 cL/L also. In general, medicinal synthetic air produced at Maxima is fit-for-purpose (fit for intended use) as it satisfies the EP requirements completely without any statistically-significant consumer's risk.

The specific producer's risk R_p^* relevant to water vapour volume fractions greater than 67 ppm, i.e., outside the acceptance interval, is zero as such measured values are not feasible. Hence, the total specific producer's risks can differ from zero only when the measured oxygen volume fractions are outside their acceptance interval, while the water vapour volume fractions are inside their acceptance interval. In that case, the total specific producer's risk coincides with the particular specific risk for oxygen.

4.2. Risks in conformity assessment of synthetic air for CCQM-K120

The medians of the standard measurement uncertainties obtained from the RawData_CCQM.txt file were considered as the robust parameters of the NMIs' performance. Their values are for nitrogen $u(x_{1m})$ = 0.0000140 mol/mol, oxygen $u(x_{2m})$ = 0.000009 mol/mol, and for argon $u(x_{3m})$ = 0.000005 mol/mol. Corresponding total global consumer's risk R_c = 0.0079 and total global producer's risk R_p = 0.0081, calculated by Eq. (2), are practically equal.

Fig. 4

More details are shown in Fig. 4, where plot (a) demonstrates the total global consumer's risk $\boldsymbol{R_c}$ values and plot (b) the total global producer's risk $\boldsymbol{R_p}$ values, as functions of the measurement uncertainties on the ranges for nitrogen $u(x_{1m})$ = 0.000002 – 0.000235 mol/mol and oxygen $u(x_{2m})$ = 0.000002 – 0.000065 mol/mol. The measurement uncertainty for argon is kept equal to the median $u(x_{3m})$ = 0.000005 mol/mol. These ranges do not include, for simplicity, the outlying values $u(x_{1m})$ = 0.001105 mol/mol and $u(x_{2m})$ = 0.000295 mol/mol noted in Sec. 2.2.2.

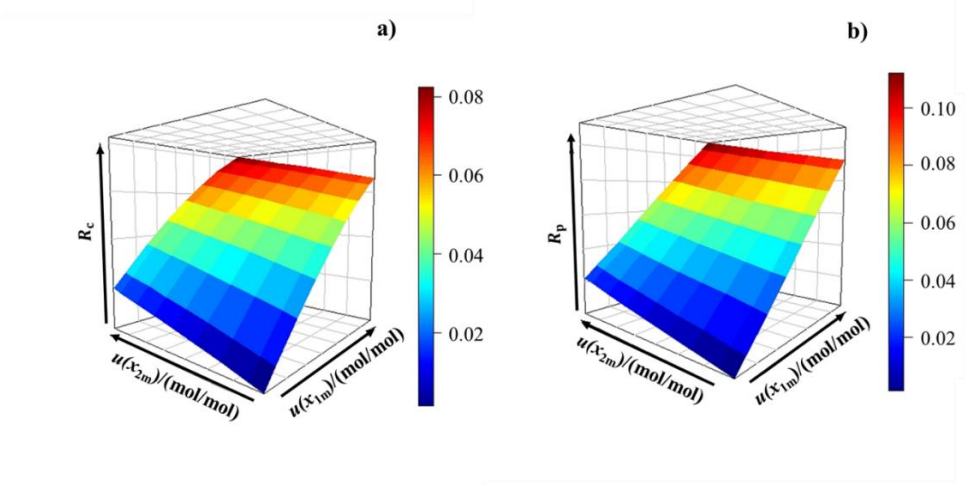

Figure 4. Dependence of total global risks on standard measurement uncertainties in synthetic air prepared in CCQM-K120. Plot (a) is the surface of the consumer's risk $\boldsymbol{R_c}$ values vs. measurement uncertainties on the ranges for nitrogen $u(x_{1m})$ = 0.000002 – 0.000235 mol/mol and oxygen $u(x_{2m})$ = 0.000002 – 0.000065 mol/mol. Plot (b) is the surface of the producer's risk $\boldsymbol{R_p}$ values vs. $u(x_{1m})$ and $u(x_{2m})$ on the same ranges as in plot (a). The colour column bars code the risk values between the minimum and maximum of the surface, each bar referring to its own plot.

Both risks increase with increasing measurement uncertainties: R_c varies from 0.0014 to 0.0824, and R_p - from 0.0015 to 0.1122. The surfaces of the risks are approximate planes twisted by

correlations. Each NMI which declared a standard measurement uncertainty in the plotted ranges, can find its own risks R_c and R_p on these surfaces.

For the calculation of total specific risks, the prior pdf and the likelihood function in Eq. (1) were approximated by relevant multivariate normal distributions according to the framework of the IUPAC/CITAC Guide [5, Eq. (34)], in order to obtain a closed (normal) expression for the corresponding multivariate posterior pdf. The approximation of a truncated normal pdf by a normal pdf is sustainable in this case study, since the distributions of the relevant quantities are very far away from the truncation limits [0, 1]. Moreover, the matrix in Table 6, rounded to three decimal digits, guarantees a proper full-rank covariance matrix for the prior pdf.

Total specific risks, evaluated by Eq. (3) for the medians of the declared standard measurement uncertainties, are plotted against the measured amount fractions of nitrogen x_{1m} and oxygen x_{2m} in Fig. 5 at the argon amount fraction equal to its measured mean $m(x_{3m})$ = 0.0093 mol/mol (Table 2).

Fig. 5

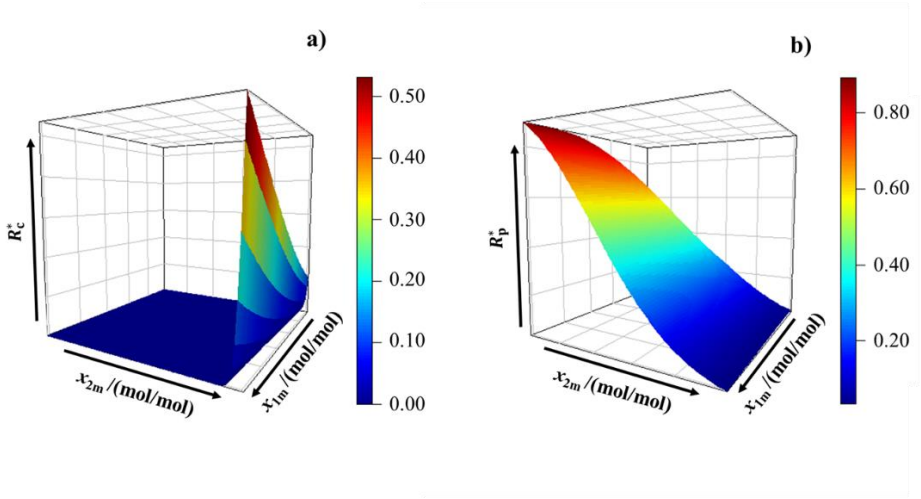

Figure 5. Dependence of the total specific risks on measured values of the amount fractions of the main components in the synthetic air for CCQM-K120. Plot (a) is the surface of consumer's risk R_c^* vs. the amount fractions of nitrogen x_{1m} from 0.7804 to 0.7809 mol/mol, and oxygen x_{2m} from 0.2094 to 0.2098 mol/mol. Plot (b) is the surface of producer's risk R_p^* vs. the amount fractions of nitrogen x_{1m} from 0.780358 to 0.780400 mol/mol, and oxygen x_{2m} from 0.209800 to 0.209827 mol/mol. The colour column bars are as in Fig. 4.

The surface of the total consumer's risk R_c^* in Fig. 5a was evaluated for intervals of x_{1m} and x_{2m} from the means $m(x_{1m})$ and $m(x_{2m})$ of the dataset (Table 2) to the tolerance limits, taking into

account the negative correlation between the two components: when values of x_{2m} increase towards $T_U(x_2)$, then x_{1m} tends to decrease. Hence, the chosen intervals were for x_{2m} from $m(x_{2m})$ to the upper tolerance limit $T_U(x_2)$, and for x_{1m} from $m(x_{1m})$ to the lower tolerance limit $T_L(x_1)$. The R_c^* values on this surface vary from 0 to 0.531. The surface of producer's risk R_p^* on Fig. 5b was calculated on the intervals of x_{1m} and x_{2m} out of their tolerance limits, taking into account again the negative correlation between the two components. The chosen intervals were for x_{2m} from the upper tolerance limit $T_U(x_2)$ to $T_U(x_2) + 3u(x_{2m})$, and for x_{1m} from the lower tolerance limit $T_L(x_1)$ to $T_L(x_1) - 3u(x_{1m})$, where $u(x_{1m})$ and $u(x_{2m})$ are the median standard measurement uncertainties. The R_p^* values vary from 0.033 to 0.894.

Note that the risks R_c^* and R_p^* specific for each NMI can be calculated at its own measured values of the air main components and associated measurement uncertainties. These risks are large in some cases, but fit-for-purpose, as the CCQM-K120 comparison was intended for evaluation of NMIs' preparative capabilities for carbon dioxide in air, whereas correspondence of the amount fractions of nitrogen, oxygen and argon to their tolerance limits was less important in this study.

5. Conclusions

A technique was developed for the evaluation of fit-for-purpose risks in conformity assessment of the chemical composition of a substance or material, based on a multivariate Bayesian approach, taking into account measurement uncertainty, correlation and the mass balance constraint.

Two datasets related to synthetic air were analysed as a case study: 1) from an industrial producer of medicinal synthetic air according to the European Pharmacopoeia, and 2) from National Metrology Institutes that participated in key comparison CCQM-K120. Each dataset included measured values with associated measurement uncertainties. The fitness for purpose of the preparation of synthetic air was reflected as total risks of false decisions on the conformity of the air composition to the tolerance limits of the contents of its main components.

The total global consumer's risk related to oxygen (or nitrogen) and water vapour volume fractions in the medicinal air were zero, whereas the calculated total global producer's risk was 0.0926. Thus, the European Pharmacopeial requirements are satisfied completely. This fitness for purpose means allowing the consumer's interests to prevail over the producer's interests.

Total global consumer's risk (0.0079) and total global producer's risk (0.0081), related to amount fractions of nitrogen, oxygen and argon in air prepared for the CCQM-K120 comparison at the median standard measurement uncertainties of the NMIs, are practically equal. These risks are also fit for purpose, as the comparison was intended for evaluation of NMIs' preparative capabilities for carbon dioxide in air, and close correspondence of the amount fractions of the air main components to their tolerance limits was less important here.

Also, it has been shown how total risks, specific for a batch of the medicinal air or for one of the NMIs that took part in the CCQM-K120 comparison, depend on measured values and their associated measurement uncertainties.

CRediT authorship contribution statement

Francesca R. Pennecchi: Methodology, Formal analysis, Software, Visualization, Writing – Reviewing and editing; **Ilya Kuselman**: Conceptualization, Project administration, Visualization, Writing - Original draft preparation; **D. Brynn Hibbert**: Writing – Reviewing and editing; **Michela Sega**: Validation; **Francesca Rolle**: Resources; **Vladimir Altshul**: Resources.

Declaration of 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.

Acknowledgements

This research was supported in part by the International Union of Pure and Applied Chemistry (Project 2019-012-1-500).

References

[1] ISO/IEC 17000, Conformity assessment. Vocabulary and general principles, 2020.

[2] JCGM 100, Evaluation of measurement data – Guide to the expression of uncertainty in measurement. http://www.bipm.org/en/publications/guides/, 2008 (accessed 20 February 2021).

[3] S.L.R. Ellison and A. Williams (Eds.), Eurachem/CITAC Guide: Quantifying uncertainty in analytical measurement. https://www.eurachem.org/index.php/publications/guides/, 2012 (accessed 20 February 2021).

[4] JCGM 106, Evaluation of Measurement Data – The Role of Measurement Uncertainty in Conformity Assessment. http://www.bipm.org/en/publications/guides/, 2012 (accessed 20 February 2021).

[5] 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 (IUPAC Technical Report). *Pure Appl Chem*, 93 (2021) 113-154. http://dx.doi.org/10.1515/pac-2019-0906.

[6] A.M.H. van der Veen, K. Hafner, Atomic weights in gas analysis, *Metrologia* 51 (2014) 80-86. https://doi.org/10.1088/0026-1394/51/1/80.

[7] M.J.T. Milton, G.M. Vargha, A.S. Brown, Gravimetric methods for preparation of standard gas mixtures, *Metrologia* 48 (2011) R1-R9. https://goi.org/10.1088/0026-1394/48/5/R01.

[8] F.R. Pennecchi, A. Di Rocco, I. Kuselman, D.B. Hibbert, M. Sega, Correlation of test results and influence of a mass balance constraint on risks in conformity assessment of a substance or material. *Measurement* 163 (2020) 107947. http://dx.doi.org/10.1016/j.measurement.2020.107947.

[9] F.R. Pennecchi, I. Kuselman, A. Di Rocco, D.B. Hibbert, A. Sobina, E. Sobina, Specific risks of false decisions in conformity assessment of a substance or material with a mass balance constraint – A case study of potassium iodate. *Measurement* 173 (2021) 108662. https://doi.org/10.1016/j.measurement.2020.108662.

[10] F.R. Pennecchi, I. Kuselman, A. Di Rocco, D.B. Hibbert, A.A. Semenova, Risks in a sausage conformity assessment due to measurement uncertainty, correlation and mass balance constraint. *Food Control* 125 (2021) 107949. https://doi.org/10.1016/j.foodcont.2021.107949.

[11] D. B. Hibbert, E.-H. Korte, U. Örnemark, Fundamental and Metrological Concepts in Analytical Chemistry (IUPAC Recommendations 2021). *Pure Appl Chem*, published online 6 August 2021, 1-56.

[12] L. Harvey, Quality Research International, *Analytic Quality Glossary* (2020) http://www.qualityresearchinternational.com/glossary/fitnessforpurpose.htm (accessed 21 February 2021).

[13] GCF/B.08/02, Guidelines for the operationalization of the fit-for-purpose accreditation approach (2014). https://www.greenclimate.fund/document/gcf-b08-02 (accessed 21 February 2021).

[14] A. Harris, Editorial: Fit for purpose: lessons in assessment and learning. *English in Education* 51/1 (2017) 5-11. https://doi.org/10.1111/eie.12134.

[15] S. Taber, N. Akdemir, L. Gorman, M. van Zanten, J.R. Frank, A "fit for purpose" framework for medical education accreditation system design. *BMC Med Educ* 20 (2020) 306. https://doi.org/10.1186/s12909-020-02122-4.

[16] S. Byrne, I. Whelan, Managing design risk through 'fit for purpose' warranties. *Johnson Winter & Slattery*, 2017. https://jws.com.au/en/insights/articles/2017-articles/managing-design-risk-through-'fit-for-purpose'-war (accessed 21 February 2021).

[17] B. Magnusson and U. Ornemark (Eds.), Eurachem Guide: The fitness for purpose of analytical methods – A laboratory guide to method validation and related topics, 2014. https://www.eurachem.org/index.php/publications/guides (accessed 21 February 2021).

[18] Analytical Methods Committee Technical Brief No. 68, Fitness for purpose: the key feature in analytical proficiency testing. *Anal Methods* 7 (2015) 7404. https://doi.org/10.1039/c5ay90052b.

[19] I. Kuselman, Design of experiment for evaluation of uncertainty from sampling in the framework of the fitness for purpose concept: a case study. *Accred Qual Assur* 13 (2008) 63–68. https://doi.org/10.1007/s00769-007-0340-z.

[20] M.H. Ramsey, S.L.R. Ellison and P. Poston (Eds.), Eurachem/EUROLAB/CITAC/Nordtest/AMC Guide: Measurement uncertainty arising from sampling: A Guide to methods and approaches, 2019. https://www.eurachem.org/index.php/publications/guides (accessed 21 February 2021).

[21] L. Pendrill, H. Karlson, N. Fischer, S. Demeyer, A. Allard, A guide to decision-making and conformity assessment - A report of the EMRP joint project NEW04 “Novel mathematical and statistical approaches to uncertainty evaluation”, 2015. https://www.researchgate.net/publication/316274532_A_guide_to_decision-making_and_conformity_assessment (accessed 22 February 2021).

[22] I. Kuselman, F.R. Pennecchi, R.J.N.B. da Silva, D.B. Hibbert, E. Anchutina, Total risk of false decision on conformity of an alloy due to measurement uncertainty and correlation of test results, *Talanta* 189 (2018) 666-674. https://doi.org/10.1016/j.talanta.2018.07.049.

[23] ISO 31000, Risk management — Principles and guidelines, 2009.

[24] Maxima Ltd, https://www.maxima.co.il/ (accessed 22 February 2021).

[25] Air, synthetic medicinal, in: European Pharmacopoeia, 10th ed., 2019, pp. 1770-1771.

[26] E. Flores, J. Viallon, T. Choteau, et al., CCQM-K120 (Carbon dioxide at background and urban level). *Metrologia* 56 (2019) 08001. https://doi.org/10.1088/0026-1394/56/1A/08001.

[27] Ellenbarrie Industrial Gases Ltd, http://ellenbarrie.com/all-gases/synthetic-air/ (accessed 22 February 2021).

[28] Air products Ltd, https://www.airproducts.co.uk/gases/calibration-mixtures#/ (accessed 22 February 2021).

[29] M. Allen and P. Edwards, Medical air: white paper. Air Liquide Healthcare, 2018. https://www.airliquidehealthcare.ca/sites/alh_ca/files/2018/07/04/medical_air_white_paper_first_edtion_15sep14.pdf (accessed 22 February 2021).

[30] Medical air. A risk assessment, Air Liquide Healthcare, 2017. https://www.airliquidehealthcare.ca/sites/alh_ca/files/2017/10/27/airliquidehealthcare-medicalair-risk-assessment.pdf (accessed 22 February 2021).

[31] Linde AG, History and technological process. Cryogenic air separation. https://www.leamericas.com/en/images/Cryogenic%20air%20separation%20brochure19_4353_tcm136-414865.pdf (accessed 11 March 2021).

[32] Servomex MimiMP 5200 Operator Manual. https://www.servomex.com/resources/product/minimp-5200/ (accessed 1 March 2021).

[33] Shaw Dewpoint Analyzer SADP Instruction Manual. https://www.shawmeters.com/wp-content/uploads/2019/08/SADP-Instruction-Manual-v2.1.1.pdf (accessed 1 March 2021).

[34] CGA/GAS - CGA G-7.1, Commodity specification for air, 2018. https://standards.globalspec.com/std/10381341/CGA%20G-7.1 (accessed 1 March 2021).

[35] JCGM 200, International Vocabulary of Metrology – Basic and General Concepts and Associated Terms, 2012. http://www.bipm.org/en/publications/guides/ (accessed 2 April 2021).

[36] ISO 6142-1, Gas analysis – Preparation of calibration gas mixtures – Par 1: gravimetric method for Class 1 mixtures, 2015.

[37] E. Flores, J. Viallon, T. Choteau, et al., International comparison, CCQM-K120a Carbon dioxide in Air at background level (380-480) µmol/mol, CCQM-K120b Carbon dioxide in Air at urban level (480-800) µmol/mol, Final report. https://www.bipm.org/utils/common/pdf/final_reports/QM/K120/CCQM-K120.pdf (accessed 11 March 2021).

[38] NIST/SEMATECH, e-Handbook of statistical methods. https://www.itl.nist.gov/div898/handbook/apr/section2/apr232.htm (accessed 21 March 2021).

[39] P. J. Huber. *Robust statistics*, New York, Chichester, Brisbane, Toronto, Singapore, John Wiley & Sons, 1981. ISBN 0-471-73577-9.

[40] W. Seidel, Mixture Models. In: M. Lovric (Ed.), *International Encyclopedia of Statistical Science*, Berlin, Heidelberg, Springer, 2011, pp. 827-829. ISBN 978-3-642-04916-3.

[41] S.P. Millard. R: Mixture of two normal distributions. https://search.r-project.org/CRAN/refmans/EnvStats/html/NormalMix.html (accessed 23 August 2021).

[42] R.B. D'Agostino, M.A. Stephens (Eds), *Goodness-of-fit techniques*, Statistics: a series of textbooks and monographs. New York, Marcel Dekker Inc, 1986. ISBN 978-0-824774875.

[43] M.D. McKay. Evaluating prediction uncertainty. Los Alamos National Laboratory - Technical Report, USA, 1995. https://doi.org/10.2172/29432.