Shades of grey in conformity assessment due to measurement uncertainty

Ilya Kuselman[1] · Francesca Pennecchi[2] · Ricardo J. N. B. da Silva[3] · D. Brynn Hibbert[4]

[1] Independent Consultant on Metrology, Israel
[2] Istituto Nazionale di Ricerca Metrologica, Italy
[3] Universidade de Lisboa, Portugal
[4] School of Chemistry, UNSW, Australia

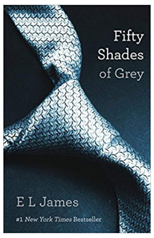
Eurachem/Eurolab/CITAC Workshop, Berlin, Germany, 19-20 Nov 2019

Outline

- Decision rules based on test results
- MU and related risks
- Grey zone and shades of grey
- Bayesian approach for risks evaluation in univariate CA
- Generalization of the approach in multivariate CA
- Examples and references

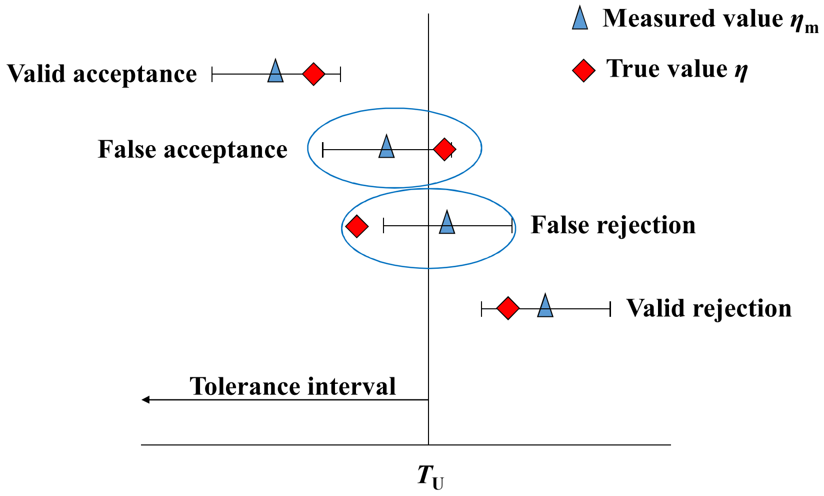
J. Coste and J. Pauchot. *A grey zone for quantitative diagnostic and screening test*. Int. J. of Epidemiology 32(2003):304-313.

M. Bataglia and D. Pewsner. *Commentary: Black and white or shades of grey?* Ibid, pp. 314-315.



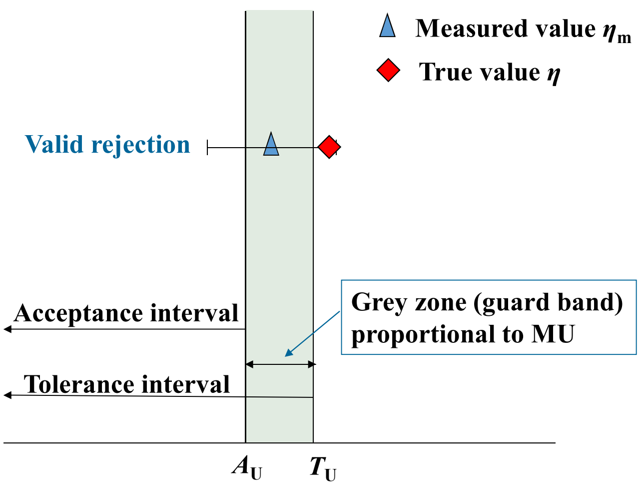
2

Measured value vs. true value



3

Acceptance interval vs. tolerance interval and grey zone



4

What do you know about risk?

5

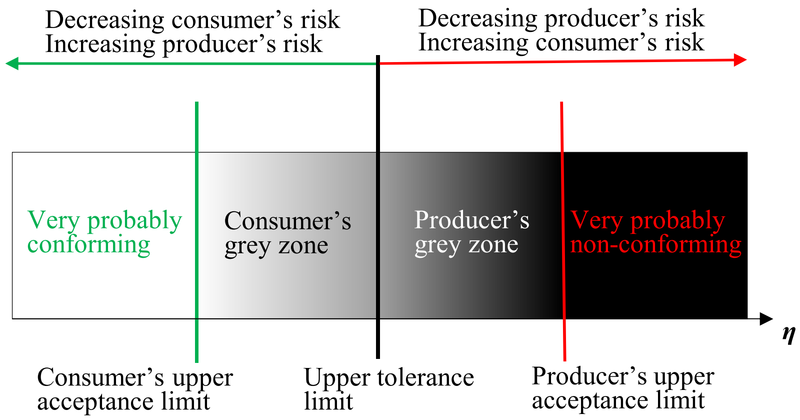
Producer's risk and consumer's risk

True product states by η and TI

Decision based on η_m and AI	Good	Bad
Pass	👍	**Consumer's risk**
Fail	**Producer's risk**	👌

6

Consumer's and producer's grey zones



7

JCGM 106:2012. Evaluation of measurement data – The role of MU in CA

JCGM 106 was transformed also into ISO/IEC Guide 98-4: 2012. It provides a methodology of CA by comparison of test results with the specification limit taking into account MU and corresponding risks, based on the Bayesian approach for univariate CA.

8

Bayesian approach

Knowledge about an item property (the measurand) can be treated as a random variable and expressed in terms of a probability density function (pdf). Such a pdf combines, by the Bayes theorem, prior knowledge of the measurand and new information acquired during the measurement:

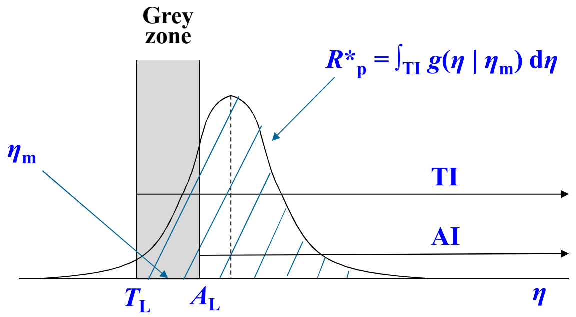
$$g(\eta \mid \eta_m) = C\, g_0(\eta)\, h(\eta_m \mid \eta),$$

where η and η_m are true and measured property values,
$g(\eta \mid \eta_m)$ is the posterior pdf,
C is the normalizing constant,
$g_0(\eta)$ is the prior pdf, and
$h(\eta_m \mid \eta)$ is the likelihood function (MU).

9

Particular specific producer's risk

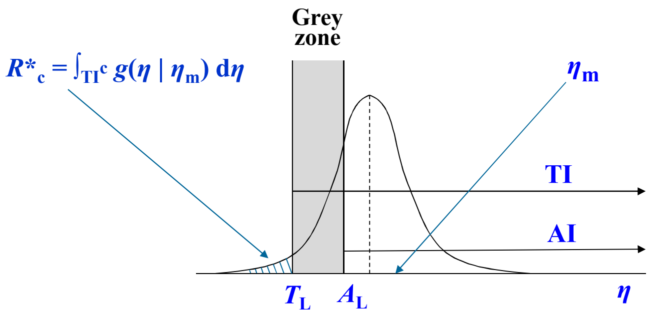
It is the probability of the false decision that a specific item should be rejected. I.e. the probability of rejecting the null hypothesis "the particular property value (component concentration or content) is in its TI", when in fact the null hypothesis is true - Type I error according to ISO/IEC 3534.



10

Particular specific consumer's risk

It is the probability of the false decision that a specific item conform. I.e. the probability of failure to reject the null hypothesis "the particular property value is in its tolerance interval", when in fact the null hypothesis is not true - Type II error.



11

Particular global producer's risk

It is the probability that a conforming property value of a particular component will be assessed as non-conforming based on a statistical analysis of earlier performed test results (the result is outside AI, but true value is within TI):

$$R_p = \int_{TI} \int_{AI^c} g_0(\eta)\, h(\eta_m \mid \eta)\, d\eta_m\, d\eta$$

This risk corresponds to the producer's risk of incorrect assessment of a particular component concentration or content in an item randomly drawn from a statistical population of such items, i.e. characterizes the material production (or objects) globally.

12

Particular global consumer's risk

It is the probability that a non-conforming property value of a particular component will be assessed as conforming based on a statistical analysis of earlier performed test results (the measured value is within AI, but true value is outside TI):

$$R_c = \int_{TI^c} \int_{AI} g_0(\eta)\, h(\eta_m \mid \eta)\, d\eta_m\, d\eta$$

The particular global consumer's risk (as the global producer's risk) corresponds to an item randomly drawn from a statistical population of such items, thus characterizes the material production or objects globally.

13

The smile of the Cheshire cat

When acceptance limit and tolerance limit coincide since the tolerance limit has already taken into account MU, the grey zone is absent. However, the same four kinds of risks, four shades of grey for each property of a material or object, are still relevant.

$$R^*_p = \int_{TI} g(\eta \mid \eta_m)\, d\eta$$

$$R^*_c = \int_{TI^c} g(\eta \mid \eta_m)\, d\eta$$

$$R_p = \int_{TI} \int_{AI^c} g_0(\eta)\, h(\eta_m \mid \eta)\, d\eta_m\, d\eta$$

$$R_c = \int_{TI^c} \int_{AI} g_0(\eta)\, h(\eta_m \mid \eta)\, d\eta_m\, d\eta$$

14

CA of a multicomponent material or object

A multicomponent material or object is not simply a mixture of n components or their sum.

When particular CA risks are acceptable, the total probability of a false decision (total consumer's risk or producer's risk) on the conformity of the material as a whole might still be significant.

The term 'total risk' is derived from the law of total probability, a fundamental rule which expresses the total probability of an outcome realized via several distinct events.

15

Total specific producer's risk

Bayes theorem for a multivariate case is in general the same, but prior, likelihood and posterior are multivariate:

$$g(\eta \mid \eta_m) = C\, g_0(\eta)\, h(\eta_m \mid \eta)$$

Therefore, total specific producer risk is the probability that true property values of all components in a specific rejected item are conforming, when at least one of the measured property values is outside its AI:

$$R^*_{total(p)} = \int \cdots \int_{TI} g(\eta \mid \eta_m)\, d\eta, \quad TI = [TI_1 \times TI_2 \times \ldots \times TI_n]$$

16

Total specific consumer's risk

It is the probability that a specific accepted item does not conform as a whole, when the true property value of at least one component is outside its TI:

$$R^*_{\text{total(c)}} = 1 - \int \cdots \int_{\text{TI}} g(\eta \mid \eta_{\text{m}})\, d\eta$$

In other words, the 'total specific consumer's risk' is the probability of a false decision that an item is conforming, based on the measured values not exceeding their upper acceptance limits. For example, $c_{im} \leq A_{Ui}$ for all i, while one true value c_j exceeds its upper tolerance limit ($c_j > T_{Uj}$).

17

Total global producer's risk

It is the probability that an item with conforming property true values of all the components will be rejected based on a statistical analysis of earlier performed measurement (chemical analytical test) results.

For simplicity, for the case of two components under control ($n = 2$):

$$R_{\text{total(p)}} = P(\bar{C} \cap \bar{B}) = P(\bar{B} \cap \bar{C}_1) + P(\bar{B} \cap \bar{C}_2) - P(\bar{B} \cap \bar{C}_1 \cap \bar{C}_2),$$

where $P(\bar{B} \cap \bar{C}_1 \cap \bar{C}_2) =$

$$= \int_{\text{AI}_1{}^c} \int_{\text{AI}_2{}^c} \int_{\text{TI}_1} \int_{\text{TI}_2} g_0(\eta_1, \eta_2)\, h(\eta_{1\text{m}}, \eta_{2\text{m}} / \eta_1, \eta_2)\, d\eta_1\, d\eta_2\, d\eta_{1\text{m}}\, d\eta_{1\text{m}}$$

18

Total global consumer's risk

It is the probability that an item with non-conforming property true values of one or more components will be accepted based on a statistical analysis of earlier performed measurement results.

$$R_{total(c)} = P(C \cap B) = P(C \cap B_1) + P(C \cap B_2) - P(C \cap B_1 \cap B_2),$$

where $P(C \cap B_1 \cap B_2) =$

$$= \int_{AI_1} \int_{AI_2} \int_{TI_1^c} \int_{TI_2^c} g_0(\eta_1, \eta_2)\, h(\eta_{1m}, \eta_{2m} \mid \eta_1, \eta_2)\, d\eta_1\, d\eta_2\, d\eta_{1m}\, d\eta_{1m}$$

Similar expressions for more than two components under control can be formulated using the same multivariate modeling.

19

How many shades of grey are in CA?

There are four kinds of the particular risks for each *i*-th property value (component concentration or content) of a material, and four kinds of the total risks.

For $n > 1$ components under control one can distinguish $4n$ particular risks and 4 total risks, altogether $4(n + 1)$ kinds of risks of false decisions – shades of grey.

For example, for two components this means 4(2 +1) = 12, for three components – 16, and for four components – 20 kinds of risks as shades of grey.

20

Conclusions

- For each particular property of a material or object there are four kinds of risks in CA - four “shades of grey”.
- When particular CA for each component of a material or object is successful, the total risk, i.e. the total probability of a false decision concerning conformity of the material as a whole may still be significant.
- For n >1 components under control one can distinguish $4n$ particular risks and 4 total risks, altogether $4(n+1)$ kinds of risks of false decisions as shades of grey.

21

Conformity assessment

ISO/CASCO: 104 PM representing their countries, 35 OM, 31 published documents, 6 under development.

ISO/IEC 17000:2004 - Conformity assessment. Vocabulary and general principles: conformity assessment (CA) is demonstration that specified requirements relating to a product, process, system, person or body are fulfilled.

https://www.youtube.com/watch?v=_Sjp58GUJwk&app=desktop

Conformity of a product (e.g. a material) is assessed before it is placed on the market. The CA procedure for each product is specified in the applicable product legislation.

22

Comparison of a result with the tolerance limit

Comparing a measurement/test result with the specification, regulation or legal (tolerance) limits of the material, one should decide whether the tested concentration conforms or not.

23

Guidelines for setting tolerance limits

Within tolerance limits a product, material or object is expected to perform its stated and intended function for consumer's use or to be acceptable from medical, forensic or another point of view.

ILAC G8. *Guidance on Decision Rules and Conformity with Requirements* (2019).

Eurolab Technical Report No. 01/2017. *Decision rules applied to conformity assessment* (2017).

WADA Technical Document TD2019DL. *Decision limits for confirmatory quantification of threshold substances*.

24

Acceptance interval and limit(s)

AI is interval of permissible measured values of a property value of a component, taken into account MU.

Acceptance limit is upper or lower bound of an AI.

S.L.R. Ellison and A. Williams (eds). EURACHEM/CITAC Guide: *Use of Uncertainty Information in Compliance Assessment* (2007)

S. Puydarrieux et al. *Role of measurement uncertainty in conformity assessment*. 19th Int. Congress of Metrology (2019), https://doi.org/10.1051/metrology/201916003

25

ILAC guidelines on decision rule risks

ISO/IEC 17025:2017 recognizes that no single decision rule can cover all situations applicable to statements of conformity.

ILAC G8:2019 aims to provide an overview of decision rules in conformity. It doesn't enter into details regarding statistics and mathematics but refers readers to the relevant literature.

This means that some labs, their personnel and customers will be required to improve their knowledge related to decision rule risks and associated statistics.

26

Computational details

Calculation of parameters of the posterior multivariate normal distributions and corresponding specific risk values were performed in the R programming environment. Simulation of the posterior distribution is also possible by Markov Chain Monte Carlo (MCMC) method, using the Metropolis-Hasting algorithm with MS Excel.

Global risks evaluated also using R, as well as Monte Carlo (MC) simulation and Cholesky decomposition of the covariance matrix with MS Excel, produced satisfactorily close calculation results.

27

More about computation and all the work

Core of the R codes for calculations of the risks for uncorrelated and correlated data is available in our papers in Chemosphere 202(2018):165-176, and Talanta (2018)189: 666-674, respectively.

User-friendly MS Excel spreadsheets for the same purposes are described in the papers in Chemometrics and Intelligent Laboratory Systems 182(2018):109-116 and 188(2019):1-5.

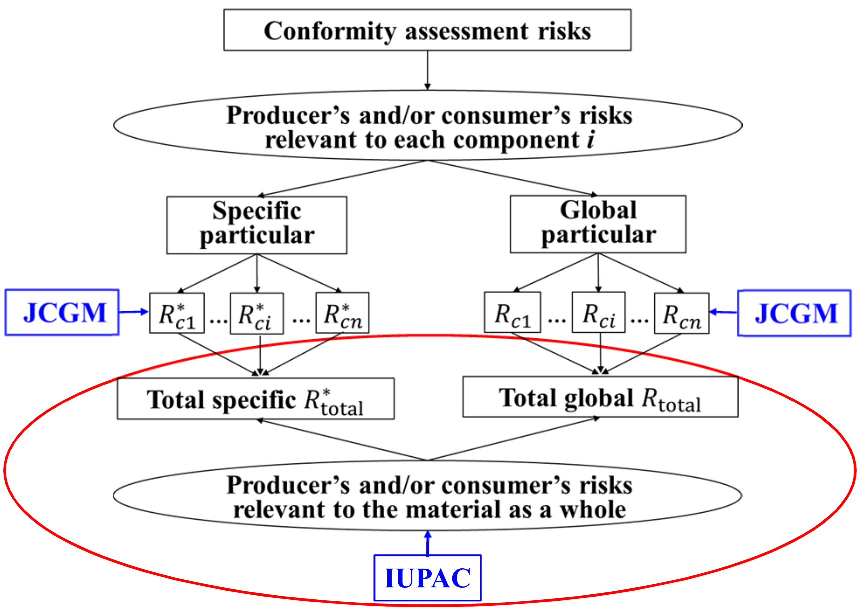
Publication of *IUPAC/CITAC Guide for evaluation of risks of false decisions in CA of a multicomponent material or object due to MU (IUPAC Technical Report)* is planned at beginning 2020.

28

Examples

1. Conformity assessment of denatured alcohols for customs control – Talanta 164(2017):189-195
2. Conformity assessment of concentration of total suspended particulate matter in ambient air – Chemosphere 202(2018):165-176
3. Conformity assessment of a medication with correlated data - Talanta 174(2017):789-796
4. Conformity assessment of chemical composition of a PtRh alloy - Talanta 189(2018):666-674

29



30