The next GUM and its proposals: a comparison study

J.C. Damasceno [1,2], P.R.G. Couto [2,3]

[1] Divisão de Metrologia de Materiais – Inmetro; [2] Núcleo de Incerteza de Medição e Estatística Aplicada (Nimea) – Inmetro; [3] Divisão de Metrologia Mecânica – Inmetro

E-mail: jcdamasceno@inmetro.gov.br

Abstract: The Guide to the Expression of Uncertainty in Measurement (GUM) is currently under revision. New proposals for its implementation were circulated in the form of a draft document. Two of the main changes are explored in this work using a Brinell hardness model example. Changes in the evaluation of uncertainty for repeated indications and in the construction of coverage intervals are compared with the classic GUM and with Monte Carlo simulation method.

Keywords: Measurement uncertainty; GUM; Monte Carlo; Brinell hardness.

1. INTRODUCTION

A few decades ago, no universal guidance for the expression of measurement uncertainty existed. In this way, metrology laboratories in the world could not compare their results with good reliability. Since 1993, when the first edition of GUM (Guide to the Expression of Uncertainty in Measurement) was published, this scenario has changed. The GUM was really well received by metrology laboratories as it harmonized the way measurement uncertainty was evaluated.

The GUM received some minor revisions later in 1995 and again in 2008, when it was made free of charge as a JCGM (Joint Committee for Guides in Metrology) document [1]. GUM supplements [2-5] were also published since then to extend and generalize the applicability of the guide without changing the main document. However, the extension of the GUM concepts in its supplements caused an unavoidable inconsistency with the former GUM guide. The supplements incorporated a number of newer concepts, based on Bayesian analysis and numerical methods that do not coexist well with concepts exposed in the original GUM.

JCGM then proposed a new revision of the GUM and a first draft was circulated among national institutes of metrology in 2012. Remarkable changes were made that could affect the way laboratories deal with the measurement uncertainty results. This revision is still being discussed and some information about it has also been released elsewhere [6].

In this work, we discuss some of these changes and use an example to show their consequences.

2. THE GUM AND ITS NEW PROPOSALS

Two main proposals of the new GUM draft are discussed in this work: a correction in the uncertainty for repeated measurement indications and the construction of coverage intervals for the final measurand result.

2.1 Uncertainty for repeated indications

In the classic GUM, repeated indications are treated like having a normal distribution and the uncertainty u_R associated with it is evaluated by:

$$u_R = \frac{s}{\sqrt{n}} \tag{1}$$

where s is the standard deviation of the indicated values and n is the number of repeated indications. This evaluation is not consistent with the GUM supplement 1 [2], where repeated indications are treated as Student's t distributions to account for the lack of degrees of freedom, or the low number of indications. The new proposal for the GUM is to consider repeated indications as t distributions, just like in supplement 1. So its uncertainty would be evaluated as:

$$u_R = \left(\frac{n-1}{n-3}\right)^{1/2} \frac{s}{\sqrt{n}} \tag{2}$$

2.2 Coverage intervals

Once the combined uncertainty is obtained after propagation, the classic GUM uses a t distribution to construct a coverage interval for a designed coverage probability. It uses the Welch–Satterthwaite equation to evaluate the effective degrees of freedom for the final distribution and then calculates a coverage factor k to expand the combined uncertainty u, thus:

$$U = ku \tag{3}$$

where U is the expanded uncertainty.

The draft for the new GUM proposal suggests that the final coverage interval cannot be reliably determined if only an expectation y and a standard deviation $u(y)$ is known, mainly if the final distribution deviates significantly from a normal or t distribution. Thus, they propose distribution-free coverage intervals in the form $y \pm U_p$, with $U_p = k_p u(y)$: (a) If no information is known about the final distribution, then a coverage interval for the measurand Y for coverage probability of at least p is determined using $k_p = 1/(1-p)^{1/2}$. If $p = 0,95$, a coverage interval of $y \pm 4,47u(y)$ is evaluated. (b) If it is known that the distribution is unimodal and symmetric about y, then $k_p = 2/[3(1-p)^{1/2}]$ and the coverage interval $y \pm 2,98u(y)$ would correspond to a coverage probability of at least $p = 0,95$.

3. CASE STUDY: BRINELL HARDNESS

A Brinell hardness test is used in this work as a case study for the application and discussion of the GUM and its new proposals. In this example, a sphere made from a very hard material is loaded onto the surface of the material sample. This procedure leaves an indentation mark with a diameter that is inversely proportional to the hardness of the material. The Brinell hardness BH is calculated in [kgf/mm²] as:

$$BH = \frac{2F}{\pi D(D - \sqrt{D^2 - d^2})} \tag{4}$$

where F is the applied load in [kgf], D is the diameter of the sphere in [mm] and d is the diameter of the indentation mark in [mm].

In this particular case study, there are three input quantities, each of them with one source of uncertainty: (a) uncertainty of F is given by a certificate as 2%, with $k = 2$; (b) uncertainty of D is certified as 0,01 mm, with $k = 2$; and (c) uncertainty of d is given by the repeatability of 5 measurements.

In addition to the GUM and its new proposals, supplement 1 was also employed to further compare the results. Monte Carlo numerical simulations were performed using 200.000 iterations.

4. RESULTS AND DISCUSSION

Table 1 summarizes the values for the input quantities used for the Brinell hardness example.

Table 1. Input quantities for the Brinell hardness example.

Quantity	Value
Load (F)	3000 kgf
Sphere diameter (D)	10 mm
Indenter mark diameter (d)	3 mm

The standard uncertainties for the quantities were calculated as: $u(F) = 0,02 * 3000/2 = 30$ kgf, for the load; and $u(D) = 0,01/2 = 0,005$ mm, for the sphere diameter. The indenter mark measurement was repeated 5 times and the average was found as 3,00 mm, with a standard deviation of 0,0791 mm. The standard uncertainty for d was calculated using the classic GUM approach (equation 1), as well as using the new GUM proposition (equation 2).

In order to check the impact of the sample size of the repeated indications (n) on the uncertainty propagation using both methods, all calculations were performed using n as 5, 10, 15, 20, 30 and 40. Results for the resulting combined uncertainty are shown on Figure 1.
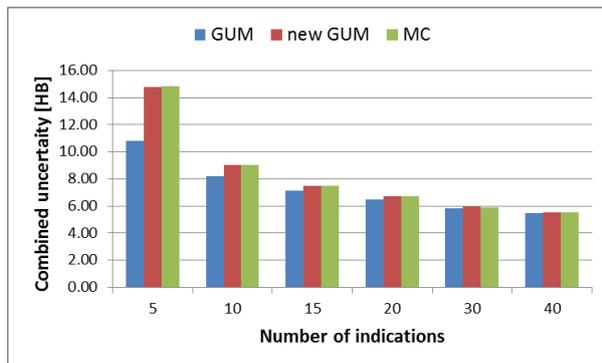
Figure 1. Results for the GUM, new GUM and Monte Carlo approaches for the combined uncertainty of Brinell hardness tests.

As can be noted, for all sample sizes the combined uncertainty obtained for the new GUM approach was very similar to the one obtained with the Monte Carlo simulation. On the other hand, for small sample sizes, the classic GUM approach result in underestimated combined

uncertainties in comparison with the other two methods. This indicates that the new GUM proposal is really a better option then the classic GUM, particularly for small sample sizes, considering that the Monte Carlo simulation is the most realistic evaluation as it is free from many approximations due to the law of propagation of uncertainties.

Table 2 shows the comparison for the expanded uncertainties. In this case, confidence intervals were constructed as follows: (a) as per the classic GUM approach, using the effective degrees of freedom (t distribution); (b) using the corrected equation for the repeated indications (equation 2) and multiplying the combined uncertainty by $k = 1,96$ (normal distribution); (c) using the new GUM proposal of $k = 2,98$ (unimodal and symmetric distribution); (d) using $k = 4,47$ (for a completely unknown distribution); and (e) using Monte Carlo simulation. In this last case, for the sake of comparison the expanded uncertainty was calculated as the semi-interval covered by the distribution endpoints for $p = 0,95$, considering a symmetric distribution.

Table 2. Expanded uncertainty results as a function of sample size for a Brinell hardness example. All values are in HB units.

n	Classic GUM (a)	New GUM $k = 1,96$ (b)	New GUM $k = 2,98$ (c)	New GUM $k = 4,47$ (d)	MC (e)
5	27,85	28,91	43,98	65,96	28,98
10	17,39	17,70	26,93	40,40	17,93
15	14,49	14,69	22,34	33,51	14,75
20	13,03	13,17	20,04	30,06	13,21
30	11,53	11,62	17,67	26,51	11,64
40	10,74	10,81	16,44	24,65	10,86

Table 2 shows that there is a good agreement between the GUM corrected for the sample size (b) and the Monte Carlo simulation (e), for all

sample sizes. The values for the classic GUM (a) are also very close, but for small sample sizes there is a relatively small underestimation. This indicates that the use of the corrected equation (equation 2) for the t distribution is a good choice for this model. Coverage intervals for the two proposals in the new GUM (c and d) overestimated the expanded uncertainty in comparison with the Monte Carlo result.

5. CONCLUSIONS

A revision of the GUM is in progress by the JCGM. The draft of the new version presents concepts that are more consistent with the GUM supplements and includes several new proposals.

Study of an example of Brinell hardness measurement in this work showed that the use Student's t distribution for repeated indication of values in a type A uncertainty is a better option as it compensates small sample sizes in a better way than the use of the Welch–Satterthwaite formula to calculate effective degrees of freedom.

Coverage intervals proposed by the GUM draft are very conservative, raising the expanded uncertainty in least 50% in comparison with the Monte Carlo propagation.

6. REFERENCES

[1] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP e OIML, Evaluation of Measurement Data – Guide to the Expression of Uncertainty in Measurement (GUM 1995 with minor corrections). Joint Committee for Guides in Metrology, JCGM 100:2008

[2] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML, Evaluation of Measurement Data – Supplement 1 to the Guide to the Expression of Uncertainty in Measurement – Propagation of distributions using a Monte Carlo method, Joint Committee for Guides in Metrology, JCGM 101:2008

[3] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML, Evaluation of Measurement Data – Supplement 2 to the Guide to the Expression of Uncertainty in Measurement – Models with any Number of Output Quantities, Joint Committee for Guides in Metrology, JCGM 102:2011

[4] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML, Evaluation of Measurement Data – An introduction to the Guide to the Expression of Uncertainty in Measurement and related documents, Joint Committee for Guides in Metrology, JCGM 104:2009

[5] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML, Evaluation of Measurement Data – The Role of Measurement Uncertainty in Conformity Assessment, Joint Committee for Guides in Metrology, JCGM 106:2012

[6] Bich W, Cox M and Michotte C, 2016, *Metrologia* **53** S149-S159