Quality control programme for mercury human biomonitoring

Abstract

The objective of the document is to define an effective system for performing quality-control activities to ensure the reliability of mercury human biomonitoring (HBM) results. These activities are focused on the pre-analytical and analytical stages of the mercury HBM. The measures described should be seen as a general recommendation for use when planning and implementing HBM surveys at national, regional and international level. The document should be considered for use together with relevant standard operating procedures for sampling and analysis of mercury in human scalp hair, cord blood and urine.

Keywords

Contributors

Argelia Castaño Calvo (National Centre for Environmental Health, Carlos III Institute of Health, Spain)
Marta Esteban López (National Centre for Environmental Health, Carlos III Institute of Health, Spain)
Miguel Angel Lucena (National Centre for Environmental Health, Carlos III Institute of Health, Spain)

CONTENTS

Abbreviations

HBM   human biomonitoring
ID   identification

1. INTRODUCTION

Although human biomonitoring (HBM) has been widely employed in the framework of occupational exposure, it has only recently been used to assess the exposure of the general population to environmental pollutants. The extension of HBM to this field of application over the past few years has been boosted by, among others, different initiatives focused on increasing our understanding of the relationship between the environment and health.

The potential of HBM in the field of public health is an accepted fact, although the lack of harmonization between the different HBM studies/programmes can considerably limit the comparison of results, their global interpretation and subsequent translation into policy. It is, therefore, fundamentally important to develop a harmonized framework that allows the most efficient use of data obtained in HBM studies, such as in the European Union-supported projects Development of a coherent approach to human biomonitoring in Europe (ESBIO), Consortium to Perform Human Biomonitoring on a European Scale (COPHES) and its twin feasibility study DEMOCOPHES.

The organization of an HBM survey is a complex process involving professionals with different technical skills (epidemiologists, analytical chemists, toxicologists, statisticians, physicians and communications specialists), all of whom contribute to specific stages of the study. They work together to deal with the interactions between the various disciplines concerned (Fig. 1).

Fig 1. Stages of a biomonitoring study
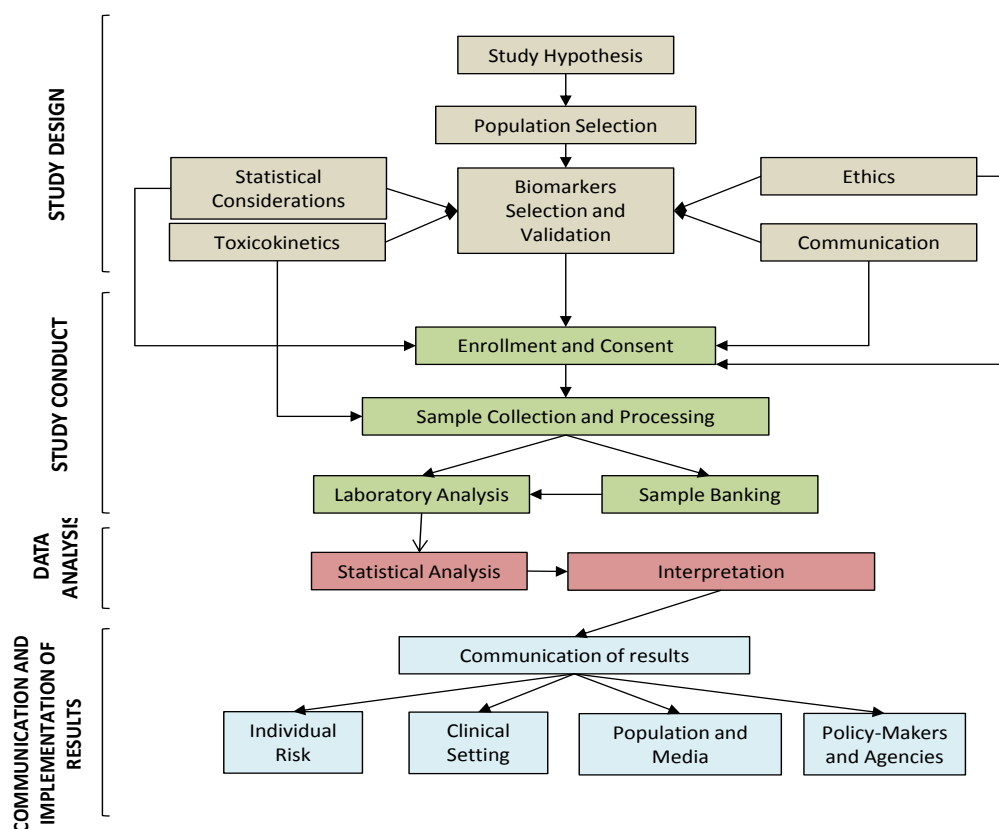
Source: National Research Council of the National Academies *(1).*

Quality control measures tend to focus on the analytical phase. All laboratories employ measures such as blanks, calibration curves and control samples during analysis in order to guarantee the reliability of their results. Control measures are, however, often missing in other stages of an HBM study that may be equally, or even more, important from a quality control standpoint.

It is important to note that all the precautions and control measures taken during chemical analysis are useless if the samples have been contaminated or altered during sampling, transport or processing. In view of this, HBM should not start in the laboratory; quality control measures must cover all steps in the pre-analytical phase, especially sampling (materials, vessels, procedures and documentation), transport (temperature, shipping requirements), sample pretreatment (centrifugation, extraction), aliquoting process and storage. Further down the line, this control should be extended to chemical analyses and subsequently to the data analysis process, by applying quality control to the databases generated.

2. QUALITY CONTROL AT THE PRE-ANALYTICAL PHASE

The pre-analytical phase has important implications for sample integrity in all studies involving biological specimens. Generally speaking, two kinds of factor can alter the sample before it is analysed.

- Influencing factors appear before the sample is collected and are specific for each biomarker. Examples of influencing factors that can modify the biomarker concentration include the biological half-life of a chemical, alcohol consumption, medication intake or individual habits such as diet. These issues should, therefore, be taken into account during the design of a study when describing the study population, statistical considerations, sampling strategies, recruitment or biomarker/matrix selection as well as during interpretation of the results.

- In the case of interfering factors, the concentration of the biomarker is modified after sampling due, for example, to external contamination, physical or chemical changes in the biomarker during transport or storage, or changes in the biological matrix such as coagulation or sedimentation. Several precautions can be taken to avoid these alterations and potential contamination during sampling and the transport, processing and storage of samples. Additionally, appropriate training of the fieldworkers is highly beneficial in this respect.

In HBM studies involving the general, presumed non-exposed population, control of the pre-analytical phase is even more important than in other kinds of study that involve biological samples due to the characteristics of such HBM studies, especially the types of substance analysed and the concentration ranges, which are usually measured in HBM. Thus, when measuring an environmental chemical there is a risk of sample contamination due to the presence of this chemical in the environment. This is particularly important in the case of ubiquitous chemicals that may even be present in the sampling material *(2)*. Additionally, as exposure to environmental chemicals occurs at low concentrations, their levels in biological matrices also tend to be low, so that the influence of potential contamination on the results is high.

It is, therefore, essential to identify and avoid possible sources of contamination, such as:

- exogenous contamination at the sampling location;

- contamination from the sampling equipment or vessels;

- contamination due to absorption of the components to be analysed into the walls of the vessel employed.

The influencing factors for the target biomarker must be identified and a sampling strategy designed to take them into account. Finally, the information required to ensure correct interpretation of the results must be recorded.

Although various tools can be employed to achieve good quality control, standard operating procedures (SOP) tend to be the most useful. An SOP is a clear, concise, comprehensive and detailed step-by-step written description of a sampling or recruitment procedure or analytical method.

SOPs can be applied at all stages of a study to provide the basic information for quality control. Their use helps different laboratories/research teams to obtain comparable results. In view of the above, SOPs for the selection of participants and the recruitment protocol should be developed together with those for sampling and the transport, processing and storage of samples in order to control, as far as posible, all the factors that can affect the sample during the pre-analytical phase.

Other control measures include the use of field blanks during fieldwork or the collection of replicate samples. Different kinds of blank can be used during fieldwork to assess potential contamination of a sample during sampling or transport until its arrival at the laboratory. An empty vessel or tube (from the same batch as the rest of the material) can be considered a collection blank. These blanks are especially useful when prescreening of the material has not been carried out and for identifying environmental contamination. Blanks can also be prepared during the aliquoting process. Blanks should be treated and manipulated as though they were real samples in order to evaluate potential contamination during the real process.

It is also essential to ensure that the sample collected is representative and reflects the composition of the original. Thus, during the aliquoting process, all samples must be homogenized before being divided. Checklists containing the necessary materials or important points to be checked are also good control tools.

Special attention must be paid to the sampling and storage materials as different kinds of interference have been described between the materials from which the vessels or tubes where samples are collected were made and the target chemical. For example, glass must be avoided when metals are analysed *(3)*. Likewise, some types of plastic can increase biomarker concentration values, for example, in the case of bisphenol A or phthalates *(2)*.

Control of the sampling and storage materials is crucial in HBM because, as noted above, the concentrations measured are usually in the range of parts per million or parts per billion (or even lower), meaning that a minimal background contamination can have important consequences for the final results. In order to control this potential source of errors, the following approaches can be designed to control the sampling material.

- Prescreening of the sampling material, consisting of screening a batch of collection tubes or vessels prior to sampling, should be undertaken to ensure that the background contamination is negligible (that is, < LOD (Limit of Detection)) and will not contribute to the final measurement. This precaution should be extended to the storage material.

- In some situations, the sampling and storage materials can be precleaned to eliminate any potential background contamination. For example, vessels employed to collect urine for metals analysis can be washed with a dilute nitric acid solution to eliminate any traces of metal from them. The effect of such a pretreatment should be checked by analysing 5% of the pretreated material.

- Materials certified to contain less than a minimum concentration of the target biomarker can be used. Some commercial materials are provided with certificates indicating the absence or minimum content of a specific chemical. For example, special tubes are available for trace metals analysis in blood samples.

The sampling time is a highly critical point during the pre-analytical phase. Correct sample collection requires an SOP containing detailed step-by-step instructions. Likewise, a written record of every event that occurs during sampling and all sample-related parameters (date and time of collection, volume, length and colour) are other useful quality control measures. Such steps can help to identify, for example, cross-contamination of a sample (for example, a urine sample contaminated with blood due to maceration caused by delivery). Additionally, well-documented fieldwork facilitates communication and helps to avoid misunderstandings and errors in the fieldworkers' team and between fieldworkers and laboratory staff.

Finally, sample traceability must be guaranteed, necessitating unambiguous identification of the specimens and related documents (questionnaires, personal data).

The quality of the labels used should, therefore, be checked to ensure that the identification (ID) code remains legible irrespective of temperature and humidity and, of course, that the label remains stuck to the tube, vessel or document.

After sampling, the samples collected should be transported under the conditions required to maintain their integrity. This is another critical control point. Transport to the laboratory must be done in compliance with the shipping regulations for biological materials.

The final step in the pre-analytical phase is sample storage and biobanking (if such is planned), although a previous step, namely reception of the samples and the acceptance/rejection criteria, should not be overlooked. Although these aspects are sometimes forgotten, they are crucial control points.

When samples arrive at the laboratory, the integrity of the packaging and the conditions of the sample tubes and vessels should be checked. Any problem encountered, such as broken or damaged packaging or a spilled sample, must be documented. To ensure that this is carried out correctly, it is advisable to establish a sample reception protocol that specifies the items to be checked and allows problems to be recorded on a sample registration sheet (Annexes 1 and 2).

The requirements for sample transport should be defined beforehand so as to establish the critical points to be checked. Checking should be performed from the outside in, that is, first verifying the state of the packaging and then opening it and continuing the process. If the biological specimens are accompanied by questionnaires or other documents, these should also be checked during the reception of the simple, when the previously defined sample acceptance/rejection criteria should be applied.

Although SOPs are essential support tools, they are not a comprehensive solution and should be complemented by trained laboratory staff and fieldworkers.

3.   QUALITY CONTROL AT AN ANALYTICAL PHASE

From an analytical point of view, it is essential to establish a quality assurance/quality control programme to ensure the reliability and comparability of results. Such programmes should cover both the basic quality assurance/quality control measures routinely applied in analytical laboratories as well as external action to ensure the comparability and quality of the results.

Internal quality controls are a basic tool in analytical laboratories as it is essential to harmonize control activities as well as the SOPs used to obtain laboratory results. Quality control activities must, therefore, be one of the fundamental points described in the working procedure, and tolerance criteria must be well-established before any assays are performed.

Refinement of the method must take into account that blank controls, repeatability controls, reproducibility controls or veracity are parameters that must be considered when evaluating the performance of the method. No external results should be reported in the absence of correct results from internal quality controls associated with the analysis and confirmation from the laboratory of compliance with these requirements.

The present procedure is concerned with the performance of quality controls associated with instrumental methods, typically methods based on the preparation of working curves onto which the results of test samples can be interpolated.

Interlaboratory comparison can be seen as a measure of the capacity of a laboratory. In order to obtain enough information about the performance of a laboratory, at least three rounds must be considered: before, during and after analysis of the study samples. In this way, the accuracy of the results of the participants can be evaluated and the validity of the study ensured.

Participation in each of these rounds must be evaluated according to the defined criteria. Unsatisfactory results in some of the rounds must be investigated and the possible causes of malfunction must be eliminated and corrected. In this respect, interlaboratory exercises can be used to demonstrate the adequate performance of laboratories in comparison with others.

3.1. Internal quality controls

3.1.1. Standards

Internal quality controls should be performed using certificated standards, when these are available. Such standards must be certified confirming their traceability to international standards. In addition, they must have an associated uncertainty in order to evaluate the confidence intervals and allow the laboratory to determine the accuracy of its results.

Any manipulation of these standards (such as dilution in order to obtain lower concentrations of the nominal value) means that the laboratory must calculate the new uncertainty based on the initial uncertainty of the standard and all contributions associated with the volumetric equipment used during preparation. If the analytical method has been suitably validated, these uncertainties will have been considered and consequently included in the defined tolerances of the validation.

3.1.2. Equipment

Equipment that may affect the result of the assay must be calibrated. In this regard, laboratories should have previous predefined tolerances that can be used to accept or reject the results of these calibrations.

Volumetric equipment must achieve the tolerance established for its class although, as a general rule, only class A volumetric material should be used.

Precision balances must be used to weigh standards or samples, when required. For example, in the case of hair analysis, no less than 30 mg of the sample or standard should be weighed if the balance has a resolution of 0.1 mg (four decimal point balance). If the laboratory has a five decimal point balance, it should be used to weigh no less than 3 mg. Weight measurements below these values introduce errors that could affect the uncertainty in the final result or increase the error in the analysis.

3.1.3. Sample conservation

Sample conservation is critical in obtaining valid results. The laboratory must have written procedures to prevent sample degradation or contamination. Storage conditions (temperature, luminosity, air- and water-tightness, humidity and storage time) should be defined.

Urine or blood samples must be stored refrigerated (<5 °C) in the dark, in an air- and water-tight container, prior to analysis. Hair samples can be stored at room temperature but must be kept away from moisture.

Laboratories must ensure the impossibility of sample contamination. Sample manipulation must be carried out in clean areas. Blank determinations, which should be treated the same as samples, can give a good indication of the cleanliness of the process.

3.1.4. Preparation of calibration curves

The laboratory must prepare a calibration curve at least every three months and use it until the next one is prepared. Curves with a minimum of five points must be prepared. The range of the curve should cover the expected values for all samples, or at least the vast majority of them.

If the laboratory has a method validation in which parameters associated with the curve have been obtained, with tolerances for these parameters, the working calibration curve must comply with the acceptance criteria obtained.

If no method validation is available, the laboratory must establish beforehand acceptance criteria for at least two of the parameters regression coefficient, linearity coefficient and slope.

Regression coefficient

The regression coefficient *(r)* is a way of determining the proportion of the total variability of a dependent variable (y) in relation to its average that is explained by the regression model. This parameter is a good measure of the goodness-of-fit of the regression curve.

The regression coefficient is considered to be adequate if it is higher than that indicated in Table 1 for the permitted confidence level and with the corresponding degrees of freedom. Usually, a 95% confidence interval is accepted as being appropriate. This interval corresponds to the column with the value 0.05. N is the number of points that have been used to construct the curve.

Table 1. Critical values for Pearson's r in a unilateral test
according to the degrees of freedom (N-2)

N - 2	0.05	0.025	0.01	0.005
1	0.988	0.997	0.9995	0.9999
2	0.900	0.950	0.980	0.990
3	0.805	0.878	0.934	0.959
4	0.729	0.811	0.882	0.917
5	0.669	0.754	0.833	0.874
6	0.622	0.707	0.789	0.834
7	0.582	0.666	0.750	0.798
8	0.549	0.632	0.716	0.765
9	0.521	0.602	0.685	0.735
10	0.497	0.576	0.658	0.708
11	0.476	0.553	0.634	0.684
12	0.458	0.532	0.612	0.661
13	0.441	0.514	0.592	0.641
14	0.426	0.497	0.574	0.623
15	0.412	0.482	0.558	0.606
16	0.400	0.468	0.542	0.590
17	0.389	0.456	0.528	0.575
18	0.378	0.444	0.516	0.561
19	0.369	0.433	0.503	0.549
20	0.360	0.423	0.492	0.537
21	0.352	0.413	0.482	0.526
22	0.344	0.404	0.472	0.515
23	0.337	0.396	0.462	0.505
24	0.330	0.388	0.453	0.496
25	0.323	0.381	0.445	0.487
26	0.317	0.374	0.437	0.479
27	0.311	0.367	0.430	0.471
28	0.306	0.361	0.423	0.463
29	0.301	0.355	0.416	0.456
30	0.296	0.349	0.409	0.449
35	0.275	0.325	0.381	0.418
40	0.257	0.304	0.358	0.393
45	0.243	0.288	0.338	0.372
50	0.231	0.273	0.322	0.354
60	0.211	0.250	0.295	0.325
70	0.195	0.232	0.274	0.302
80	0.183	0.217	0.256	0.283
90	0.173	0.205	0.242	0.267
100	0.164	0.195	0.230	0.254

Source: http://www.uv.es/meliajl/Docencia/Tablas/TablaR.PDF *(4).*

Linearity coefficient

The linearity coefficient *(Cm)* is a measure of the goodness-of-fit compared with a straight line

$$Cm = \left(1 - \frac{Sm}{m}\right)100$$

where
Sm is the deviation of the slope and m is the slope.

- If historic values for the linearity coefficient are available, the acceptance criteria should be established as the mean of the Cm values ($\overline{Cm}$) for the series of curves available minus $t_{Student}$ times the standard deviation for these Cm values ($SD_{\overline{Cm}}$) (lower limit) and 100 (upper limit) ($t_{Student}$ obtained from the number of values used).

$$\overline{Cm} - (t_{Student} \times SD_{\overline{Cm}}) \leq Cm_{current} \leq 100$$

- If no historic values are available, chromatographic methods must exceed a Cm value of 0.97 to be acceptable. Non-chromatographic methods must exceed a Cm value of 0.95.

Slope

The slope is the tangent of the angle of the straight line with the X axis and is a way of evaluating the sensitivity of the response obtained.

- If historic values for the slope are available, the acceptance criteria should be established as the mean of the slopes for the series of curves available minus $t_{Student}$ times the standard deviation for these slopes (lower limit) and the mean of the slopes for the series of curves available plus $t_{Student}$ times the standard deviation for these slopes (upper limit) ($t_{Student}$ obtained from the number of values used).

$$\overline{Slope} - \left(t_{Student} \; x \; SD_{\overline{Slope}}\right) \leq Slope_{current} \leq \overline{Slope} + \left(t_{Student} \; x \; SD_{\overline{Slope}}\right)$$

Quality controls for points on the calibration curve

A calibration curve can be used for a period of three months instead of a daily working curve. For daily measurements during this period, at least two points on the curve must be checked (one in the low range and the other in the high range) prior to starting the working series.

- If a method validation is available, the result obtained for these curve controls must lie between the acceptable values derived upon validation of the point concerned.

- If a slope history is available but there is no method validation, the result obtained must lie in the range

$$X_{point} \pm (t_{Student} \; x \; SD_{point})$$

   where

   X_{point} is the mean value obtained upon reading the point

   SD_{point} is the standard deviation for the values obtained upon reading the point

   $t_{Student}$ is obtained from the number of values used to obtain SD_{point}.

- If an estimation of the method reproducibility is available but there is no slope history, the result obtained must lie in the range

$$V_{point} \pm (t_{Student} \; x \; SD_{repro})$$

   where

   V_{point} is the value obtained for the control point on the curve

   SD_{repro} is the estimated standard deviation in reproducibility

   $t_{Student}$ is obtained from the number of values used to obtain SD_{repro}.

3.1.5. Analysis of test blanks

An initial test blank must be analysed prior to commencement of the daily series of assays.

If a method validation is available, the result obtained for the blank must be lower than the values obtained for the limit of detection for the method estimated during validation.

If historic blank readings are available but there is no method validation, the acceptance criterion for the blank is that the signal obtained must not exceed the mean for the series of blanks by more than three times the standard deviation for these values.

If a series of measurements for samples with a very low concentration of the analyte of interest is available but there are no historic blank readings, the acceptance criterion for the blank is that the

signal obtained must not exceed three times the standard deviation obtained for these samples of very low concentration.

If the first initial blank measurement does not meet the acceptance criteria, the system must be cleaned. A new test blank must be measured after this cleaning. This process must be repeated until an acceptable value is obtained. Once such an acceptable value has been obtained, a second blank reading must be performed to confirm the validity of the result. Consequently, if no acceptable value is obtained after measurement of the initial blank, two successive measurements that comply with the acceptance criteria must be obtained in order to be able to proceed with routine testing.

A *test blank* must be measured at least every five samples, using the same criteria as for the initial blanks.

If programming of the sample series is automatic and the results are collated at the *end of the series*, it may be necessary to increase the number of repeats of the blank (for example, three consecutive repetitions instead of just one) to ensure a correct reading.

A series of blank repetitions (for example, three) must be measured once the sample series has been completed to ensure the cleanliness of the system.

3.1.6 Duplicate samples

One out of every 10 samples must be repeated at different times during the series. If the sample is analysed in duplicate or triplicate, this repetition must consist of re-analysing in duplicate or triplicate. The results must be compared with each other.

- If a method validation is available, the results must comply with the reproducibility criteria obtained during validation.

- If no method validation is available, a compatibility index must be applied:

$$IC = \frac{|x_1 - x_2|}{\sqrt{(2SD_1)^2 + (2SD_2)^2}}$$

Source: ISO/IEC Guide 43-1:2007

where

x_1 and x_2 are the mean values obtained for each repetition of the sample; and

SD_1 and SD_2 are the standard deviations obtained for the duplicates, triplicates (or more) for each of the repetitions.

- If the sample is repeated using a single analysis rather than in duplicate, triplicate (or more), a maximum deviation for each of the samples with respect to the mean must be established (for example, a maximum of 10% of the mean value) to consider the result acceptable.

3.1.7 Quality controls

The laboratory must perform quality controls for at least four points in the test range: high, mid-point, low and limit of quantification. These test control points must differ from the control points for the curve.

One of these quality controls must be inserted, at random, every five samples to ensure that all have been sufficiently analysed as part of the laboratory's operations. In any case, and if the

values for all samples to be analysed fall within a very narrow range, the quality control can be repeated at the point closest to this range of samples.

- If a method validation is available, the results must comply with the criteria for the quality control points obtained during validation.

- If a history of results associated with the various control points is available but there is no validation, the acceptance criterion is that the value for the quality control must fall within the range

$$X_{point} \pm (t_{Student} \text{ x } SD_{point})$$

    where
    X_{point} is the mean value obtained upon reading the control point;
    SD_{point} is the standard deviation for the values obtained upon reading the control point; and
    $t_{Student}$ is obtained from the number of values used.

- If no results history is available, a maximum deviation for each of the samples with respect to the mean must be established (for example, a maximum of 10% of the mean value) to consider the result acceptable.

Blind samples

The laboratory must organize tests with blind samples at least once a year. To this end, the lead technician must prepare samples of a known concentration (but not known to the laboratory) for analysis. This blind sample must be prepared using certified standards, remnants of intercomparison samples or well-characterized samples.

- If a method validation is available, the real value for the sample must fall within the range

$$V_{SBlind} \pm I_{test}$$

    where
    V_{SBlind} is the value obtained upon analysing the blind sample; and
    I_{test} is the expanded uncertainty obtained during validation of the method.

- If reproducibility values are available for the method but there is no method validation, the real value for the sample must fall within the range

$$V_{SBlind} \pm (t_{Student} \text{ x } SD_{repro})$$

    where
    V_{SBlind} is the value obtained upon analysing the blind sample;
    SD_{Repro} is the value obtained for the standard deviation in reproducibility; and
    $t_{Student}$ is obtained from the number of values used to obtain SD_{repro}.

- If no other values are available, a maximum deviation with respect to the real value must be established in order to accept the result. This deviation can be estimated from the literature or the experience of the laboratory with similar methods or analytes.

3.2. External quality controls

Interlaboratory comparisons[1] are widely used for various purposes at national, regional and global scale. Examples of typical purposes for interlaboratory comparisons include:

---

[1] Interlaboratory comparison: organization, performance and assessment of measurements with the same or similar items by two or more laboratories according to predetermined conditions.

(i)   evaluating the performance of laboratories as regards conducting specific tests or measurements and monitoring the performance of laboratories over time;

(ii)   identifying problems in laboratories and initiating improvements which, for example, may be related to inadequate testing or measurement procedures, ineffective staff training and supervision or calibration of equipment;

(iii)   establishing the efficacy and comparability of testing or measurement methods;

(iv)   providing additional confidence to the laboratories' clients;

(v)   identifying differences between laboratories;

(vi)   instructing participating laboratories on the basis of the results of such comparisons;

(vii)   validating the stated uncertainty estimations;

(viii) evaluating the operational characteristics of a method;

(ix)   assigning values to reference materials and evaluating their suitability for use in specific testing or measurement procedures;

(x)   supporting equivalency declarations for measurements from national institutes of metrology by way of key comparisons and complementary comparisons performed on behalf of the International Bureau of Weights and Measures and associated metrological associations.

The procedures described below are mainly applicable for laboratories organizing an intercomparison study, such as reference laboratories at national level. They are also fully applicable for participants in interlaboratory comparisons.[2]

Proficiency tests[3] comprise the use of interlaboratory comparisons to determine the performance of laboratories, as indicated in points (i) to (vii). Proficiency tests are not normally concerned with purposes (viii), (ix) and (x) as it is assumed that laboratories are competent in these applications. However, they can be used to provide independent proof of the competence of a laboratory.

The steps prior to performing an intercomparison exercise are related to:

- assigning the value to the sample:

- determining the standard deviation parameter for the proficiency test, which will subsequently be needed for the calculations in the exercise;

- determining the number of repetitions to be performed by each participant; and

- confirming the validity of the sample to be analysed by way of homogeneity and stability tests. These parameters must be calculated by the organizer of the exercise.

Before sending the samples to the various participants, the organizer must prepare detailed and documented instructions.

Obviously, the organizer must analyse the quantity of samples required to conduct the exercise, taking into account the number of participants, the homogeneity and stability tests that need to be performed and the possibility of repetition, loss or damage to the sample during the transport

---

[2] Participant: laboratory, organization or person who receives the proficiency testing items and provides the results for review by the proficiency test provider.
[3] Proficiency test: assessment of the performance of participants with respect to previously established criteria by way of interlaboratory comparisons.

phase. A quantity that exceeds the strictly calculated requirements should, therefore, be considered.

3.2.1. Assigning the value to the sample

The criteria used to obtain the value against which the results submitted by the laboratories will be compared must be determined before the exercise is carried out. These criteria are as follows.

- The value for a certified reference material or a spiked sample is obtained when:

    − a sample of certified reference material is used for the proficiency test; the value of the property, and the uncertainty in this value, must be known; or

    − a raw sample spiked with quantities of the test substance is used; spiking may be performed by the organizer or by the participating laboratory using concentrated solutions supplied by the organizer.

- The result is obtained as the mean value obtained by a group of expert laboratories which have tested the sample or samples using previously accepted test methods that can be considered to be "absolute" or "reference" methods. Atypical results must be eliminated prior to calculating the mean.

- The result is obtained as the mean value calculated by the group of participating laboratories after elimination of atypical values or as the mean obtained using robust statistical methods (such as algorithm A; see below). This is a riskier system in free-access intercomparison systems as erroneous data may affect the mean, meaning that the elimination of outliers must be efficient.

3.2.2. Determination of the standard deviation for proficiency testing $\widehat{\sigma}$[4]

Various options are available for assigning the value of the standard deviation for proficiency testing.

Prescribed value

The standard deviation for proficiency testing can be assigned on the basis of compliance with standard values. This method has the advantage of best representing the purpose of the method.

Perceived value

The standard deviation for proficiency testing can be established on the basis of the prior experience of the coordinator and his/her collaborators using values obtained in the past.

When the standard deviation for proficiency testing ($\hat{\sigma}$) is obtained by prescription or perception, there is a possibility that the value selected is not realistic as regards the reproducibility of the measurement method. The following test can, therefore, be applied to ensure that the value of $\hat{\sigma}$ corresponds to the repeatability and reproducibility values obtained using the method
if
σ_R is the standard deviation of reproducibility, and
σ_R is the standard deviation of repeatability.

The interlaboratory standard deviation is calculated as:

---

[4] Standard deviation for proficiency assessment: measure of the dispersion used to evaluate the results of a proficiency test based on the information available.

$$\sigma_L = \sqrt{\sigma_R^2 - \sigma_r^2}$$

The value of the factor ϕ is subsequently calculated by substituting the values of σ_L and σ_r and the value selected for $\hat{\sigma}$ in the following equation:

$$\hat{\sigma} = \sqrt{(\emptyset \ x \ \sigma_L)^2 + \left(\frac{\sigma_r^2}{n}\right)}$$

where
n is the number of replicates that each laboratory will perform.

If the value obtained for ϕ is less than 0.5, the value selected for $\hat{\sigma}$ corresponds to a degree of reproducibility that, in practice, the laboratories will be unable to meet, in which case this value will have to be increased.

Value based on a general model
The value of the standard deviation for proficiency testing can be derived from the reproducibility value obtained for the measurement method.

For example, Horowitz has proposed the following model for evaluating the standard deviation of reproducibility using the concentration:

$$\sigma_R = 0,02 \ c^{0,8495}$$

where c is the concentration of the measure to be determined as a percentage (mass fraction).

Value based on the repeatability and reproducibility results

When the values for the standard deviations of reproducibility and repeatability are available, the standard deviation for proficiency testing can be obtained as follows:

σ_R is the standard deviation of reproducibility, and
σ_R is the standard deviation of repeatability.

The interlaboratory standard deviation is calculated as:

$$\sigma_L = \sqrt{\sigma_R^2 - \sigma_r^2}$$

The value of $\hat{\sigma}$ is calculated as:

$$\hat{\sigma} = \sqrt{\sigma_L^2 + \left(\frac{\sigma_r^2}{n}\right)}$$

Where
n is the number of replicates that each laboratory will perform.

Value based on the data obtained in a proficiency test round

The value of the standard deviation for proficiency testing can be derived from the value derived from the results reported by the participants in this round of tests. The standard deviation must be the robust standard deviation for the results reported by all participants, as calculated using algorithm A.

Order the p data in ascending order:

$x_1, x_2, \ldots, x_i, \ldots, x_p$

Order the robust means and robust standard deviations (x* and s*) for these data.

The initial values for x* and s* are:

x* = median of x_i (i = 1, 2, …, p)

s* = 1.483 median of $|x_i - x^*|$ (i = 1, 2, …, p)

Update the values of x* and s* as follows. Calculate:

δ = 1.5 s*

For each x_i (i = 1, 2, …, p) calculate:

$$x_i^* = \begin{cases} x^* - \delta, if\ x_i < x^* - \delta \\ x^* + \delta, if\ x_i > x^* + \delta \\ \quad x_i, other\ cases \end{cases}$$

Now recalculate the new values for x* and s* as:

$$x^* = \sum \frac{x_i^*}{p}$$

$$s^* = 1,134 \sqrt{\frac{\Sigma(x_i^* - x^*)^2}{(p-1)}}$$

Summing over i.

The robust estimation of x* and s* is derived from an iterative calculation until the process converges. Convergence is assumed when no changes occur between one iteration and the next in the third significant figure of the robust standard deviation and the equivalent figure of the robust mean. A computer can be programmed to carry out this method.

3.2.3. Criteria for selecting the number of measurements to be performed by each participating laboratory

Variations in method repeatability mean that biases may appear in proficiency tests. When the variation in repeatability is too high compared with the standard deviation for proficiency testing, there is a risk of obtaining unreliable results. In these circumstances, a laboratory may have a very high bias factor in one round but not in another, which would make finding the cause more difficult.

If we wish to limit the influence of variations in repeatability, the number of replicates (n) performed by each laboratory should be chosen such that:

$$\frac{\sigma_r}{\sqrt{n}} \leq 0,3\ \hat{\sigma}$$

where
σ_r is the standard deviation of repeatability established prior to the exercise (either by way of an experimental interlaboratory exercise or determined by the organizing laboratory).

If this condition is met, the standard deviation of repetition does not represent more than 10% of the standard deviation for proficiency testing.

In addition, all participating laboratories must perform the same number of replicates when participating in intercomparison tests.

3.2.4. Homogeneity test procedure[5]

When it is acceptable not to perform homogeneity tests for all measurands, a measurement method and characteristic measurand that are sufficiently sensitive to the heterogeneity of the samples will be selected.

Prepare and package the samples for a proficiency testing round, ensuring that there are sufficient samples to perform both the proficiency test and the homogeneity tests.

Select a number (g) of packaged samples at random, where $g \geq 10$. The number of samples included in the homogeneity test can be reduced if historic data for these homogeneity tests performed according to the same procedures are available.

Prepare two test portions for each sample, minimising the intratest differences as far as possible. Take these $2g$ test portions at random and perform the test for each one, completing the test series under repeatability conditions.

Calculate the mean ($\bar{x}$), intrasample standard deviation (s_w) and intersample standard deviation (s_S) as follows.

The data for a homogeneity test are represented by $x_{t,k}$
where
t represents the sample (t = 1, 2, …, g)
k represents the portion of sample (k = 1, 2).

The mean for each sample is defined as:

$$x_{t,.} = \frac{x_{t,1} + x_{t,2}}{2}$$

and the range of intertest portions as:

$$w_t = |x_{t,1} - x_{t,2}|$$

The general mean is calculated as:

$$\bar{x}_{.,.} = \sum \bar{x}_{t,.}/g$$

The standard deviation of the general mean is calculated as:

$$s_x = \sqrt{\sum(x_{t,.} - \bar{x}_{.,.})^2/(g-1)}$$

---

[5] According to International Organization for Standardization 13528:2005.

The intrasample standard deviation:

$$s_w = \sqrt{\sum w_t^2/(2g)}$$

where the summation covers all samples (t = 1, 2, …, g).

Finally, to calculate the intersample standard deviation:

$$s_S = \sqrt{s_x^2 - (s_w^2/2)}$$

Assessment criteria for the homogeneity test

Compare the inter-sample standard deviation (s_S) with the required standard deviation for proficiency testing ($\widehat{\sigma}$). The samples comply with an appropriate homogeneity criterion if:

$$s_S \leq 0{,}3\ \hat{\sigma}.$$

If this criterion is met, the intersample standard deviation does not represent more than 10% of the overall standard deviation for proficiency testing. If this criterion is not met, the coordinator may consider one of the following possibilities.

- The method used to prepare the samples to make any necessary improvements could be examined.

- A number of samples could be distributed to each participant in the intercomparison exercise in order to perform a measurement for each sample. The heterogeneity of these samples will increase the intrasample standard deviation to a value:

$$\sigma_{r1} = \sqrt{\sigma_r^2 + s_S^2}$$

  This value σ_{r1} can be used to increase the number of replicates for each participant in the exercise.

- The intersample standard deviation could be included in the standard deviation for proficiency testing, calculating $\hat{\sigma}$ as

$$\hat{\sigma} = \sqrt{\hat{\sigma}_1^2 + s_S^2}$$

  where
  $\hat{\sigma}_1$ is the standard deviation for proficiency testing without including any tolerance for sample heterogeneity.

3.2.5. Stability test procedure

The same laboratory that performs the homogeneity test must perform the stability tests. The same method and same product as in the homogeneity tests must be used.

Perform the stability tests after the homogeneity tests. The time difference between the former and the latter should be similar to the time that is estimated to pass between the preparation of samples for the intercomparison exercise and the maximum period during which the participants must present their results.

Take a number (g) of samples, where $g \geq 3$.

Prepare two test portions of each sample as described for the homogeneity tests.

Take the $2g$ portions at random to obtain a measurement result $y_{t,k}$ for each sample, performing all measurements under repeatability conditions.

Calculate the mean $\bar{y}_{..}$ of all measurements.

Assessment criteria for the stability test

Compare the mean obtained in the homogeneity test with the mean obtained in the stability test. The samples are considered to be stable if:

$$\left| \bar{x}_{..} - \bar{y}_{..} \right| \leq 0{,}3\, \hat{\sigma}$$

If this criterion is not met, sample preparation and storage must be assessed and improved if possible.

3.2.6. Instructions for participants

Prior to sending the proficiency test items, the proficiency test provider will notify participants of the expected date of arrival of the items and the date by which the results must be returned by the participating laboratory with sufficient notice.

The proficiency test provider must provide detailed and documented instructions to all participants. These instructions will include:

- the need to treat the proficiency test items in the same manner as the majority of samples tested routinely (unless the specific requirements of the programme require some deviation from this principle);
- the storage conditions;
- the test methods to be used, or allowed, when applicable;
- the procedure for preparing and conditioning the proficiency test items;
- handling instructions, including safety requirements;
- the specific environmental conditions under which proficiency testing must be performed and, if necessary, the requirement that participants must notify the pertinent environmental conditions during the measurement;
- specific instructions regarding the way in which results must be reported (such as measurement units, number of significant figures or decimal places) and instructions regarding the uncertainty in the result (if required); in the latter case, the coverage factor and, if possible, the probability of that coverage must be included;
- deadline for reporting the results;
- contact information for the provider in the event of any questions; and
- instructions regarding return of the proficiency test ítems (if applicable).

3.2.7. Calculation of statistical parameters associated with the proficiency test results

Estimation of participant bias

If *x* is the result (or mean of the results) reported by a participant for the measurement of one of the parameters to be determined in a proficiency testing round, the bias (D) can be calculated as:

$$D = x - X$$

where

X is the assigned value.

If a participant obtains a result that gives a bias higher than 3.0 $\hat{\sigma}$ or lower than -3.0 $\hat{\sigma}$, the result will be considered and marked as an "action signal". Similarly, if a participant obtains a result that gives a bias higher than 2.0 $\hat{\sigma}$ or lower than -2.0 $\hat{\sigma}$, the result will be considered and marked as a "warning signal".

A single action signal or two consecutive warning signals signify that the laboratory must start an investigation into the bias found in its results.

Percentage differences

If *x* is the result (or mean of the results) reported by a participant for the measurement of one of the parameters to be determined in a proficiency testing round, the percentage difference (D%) can be calculated as:

$$D\% = 100\ (x - X)/X$$

where
X is the assigned value.

If a participant obtains a result that gives a percentage difference higher than 300 $\hat{\sigma}$/X% or lower than -300 $\hat{\sigma}$/X%, the result will be considered and marked as an action signal. Similarly, if a participant obtains a result that gives a percentage difference higher than 200 $\hat{\sigma}$/X% or lower than -200 $\hat{\sigma}$/X%, the result will be considered and marked as a warning signal.

A single action signal or two consecutive warning signals signify that the laboratory must start an investigation into the bias found in its results.

3.2.8. z-score

The z-score value is calculated as:

$$z = \frac{(x - X)}{\hat{\sigma}}$$

where
x is the value reported by the participant
X is the assigned value, and
$\hat{\sigma}$ is the standard deviation for proficiency testing.

If a participant obtains a result that gives a z-score higher than 3.0 $\hat{\sigma}$ or lower than -3.0 $\hat{\sigma}$, the result will be considered and marked as an action signal. Similarly, if a participant obtains a result that gives a z-score higher than 2.0 $\hat{\sigma}$ or lower than -2.0 $\hat{\sigma}$, the result will be considered and marked as a warning signal.

If the proficiency test involves a small number of participants (for example, fewer than 10 laboratories), the significance of the z-score for the individual rounds must be considered with great care. In such cases, it is preferable to evaluate the combination of results from different rounds when assessing the performance of each laboratory.

3.2.9. E_n number

This parameter is calculated as:

$$E_n = \frac{x - X}{\sqrt{U_{lab}^2 + U_{ref}^2}}$$

where
X is the assigned value
U_{ref} is the expanded uncertainty in X
U_{lab} is the expanded uncertainty in x, the result obtained by the participant.

A value higher than 1.0 or lower than -1.0 is equivalent to a z-score value above or below 2.0, respectively, therefore a result of this type must be treated as defined in the z-score assessment.

3.2.10. z'-score

The z'-score value is calculated as:

$$z' = \frac{(x - X)}{\sqrt{\hat{\sigma}^2 + u_X^2}}$$

where
u_X is the (non-expanded) uncertainty of the assigned value X.

The z'-score results are interpreted in the same manner as the z-score values.

A comparison of the z-score with the z'-score shows that the z'-score values in one round may be lower than the corresponding z-score values, in accordance with a constant factor of

$$\frac{\hat{\sigma}}{\sqrt{\hat{\sigma}^2 + u_X^2}}$$

If

$$0,96 \leq \frac{\hat{\sigma}}{\sqrt{\hat{\sigma}^2 + u_X^2}} \leq 1,00$$

then the z'-score will be very close to the z-score, in which case it can be concluded that the uncertainty in the assigned value is negligible.

3.2.11. Zeta-score (ζ)

This parameter can only be used if the value assigned to the proficiency test has not been calculated using the results from the participating laboratories.

$$zeta\ score = \frac{(x - X)}{\sqrt{u_x^2 + u_X^2}}$$

where

u_x is the value of the standard uncertainty (not expanded) estimated by the participating laboratory, and u_X is the standard uncertainty (not expanded) of the assigned value X.

The ζ-score is interpreted in a similar manner as the z-score.

If successive ζ-scores higher than 3.0 are obtained, this may mean that the participant is underestimating the sources of uncertainty.

If a laboratory presents a very large bias and the uncertainty interval $X \pm U_X$ does not include the assigned value, very high values will also be obtained for the ζ-score.

3.2.12. E_Z score

The E_Z score is defined as

$$E_{Z-} = \frac{x - (X - U_X)}{U_x}$$

$$E_{Z+} = \frac{x - (X + U_X)}{U_x}$$

In these cases the expanded uncertainty is used.

- If both values (E_{Z-} and E_{Z+}) fall within the range -1.0 to 1.0, the result is considered to be satisfactory.

- If one of the two E_Z values falls within the range -1.0 to 1.0, the result is doubtful.

- If both values are less than -1.0 or greater than 1.0, the result is unsatisfactory.

4. REFERENCES AND BIBLIOGRAPHY

1. National Research Council of the National Academies. Human biomonitoring of environmental chemicals. Washington (DC): National Academies Press; 2006. https://doi.org/10.17226/11700.
2. Calafat AM, Needham LL. What additional factors beyond state-of-the-art analytical methods are needed for optimal generation and interpretation of biomonitoring data? Environ Health Perspect. 2009;117:1481–5.
3. Cornellis R, Heinzow B, Herber RFM, Christensen JM, Poulsen OM, Sabbioni E et al. Sample collection guidelines for trace elements in blood and urine. Pure Appl Chem.1995;67(8–9):1575–608.
4. http://www.uv.es/meliajl/Docencia/Tablas/TablaR.PDF.

Bibliography

International Organization for Standardization [website]. Geneva: International Organization for Standardization; 2018 (https://www.iso.org/standards-catalogue/browse-by-ics.html, accessed 31 January 2018).

ISO/IEC 17025:2005. General requirements for the competence of testing and calibration laboratories

ISO/IEC 17043:2010. Conformity assessment. General requirements for proficiency testing.

ISO 13528. Statistical methods for use in proficiency testing by interlaboratory comparisons

Annex No. SELF-EVALUATION OF LABORATORY COMPETENCE

This questionnaire aims to serve as a guideline for evaluating the proficiency of a laboratory based on existing information regarding equipment, expertise and participation in interlaboratory comparison exercises. It includes a series of questions organized into sections for collecting data from laboratories and the criteria for evaluating them.

Information concerning different aspects of the laboratory's facilities and methods of working, such as routine analyses should be collected and evaluated according to the established criteria.

Laboratory evaluation questionnaire

GENERAL INFORMATION

1. Data for the person completing the questionnaire.

    Name:
    ....

    Position:
    ....

    Company:
    ....

    Address:
    ....

    City:
    ....

    Post code:
    ....

    Country:
    ....

    email:
    ....

    Telephone:
    ....

2. Which analyses do you perform in your laboratory?

    ☐ Mercury in hair   ☐ Mercury in urine   ☐ Mercury in cord blood

3. Which analytical technique(s) do you use?
    ....

4. Please specify the type, manufacturer and model of your analytical apparatus
    ....

5. What is the minimum amount of scalp hair/urine/blood required for measurements?
    .... mg or mL

METHOD INFORMATION

6. Is the analytical procedure for mercury in scalp hair/urine/cord blood accredited?

26

☐ No   ☐ Yes

If yes, please include your Technical Annex number:
……………………………………………………………………………………………………

7.   Do you have a general standard operating procedure for the validation of analytical methods?

☐ No   ☐ Yes

8.   Is there a standard operating procedure for the analysis of mercury in ………… in your laboratory?

☐ No   ☐ Yes

9.   Do you have a validated method for the analysis of mercury in .…………?

☐ No   ☐ Yes

10.   Please complete the following information about your analytical method.

Interseries repeatability   ………………%

Limit of quantification   ………………

Limit of detection   ………………

Accuracy   ………………

Uncertainty   ………………

11.   How do you calculate your interseries repeatability?

………

12.   How do you calculate your limit of quantification?

………

13.   How do you calculate your limit of detection?

………

14.   How do you calculate your accuracy?

………

15.   What components do you use to calculate your uncertainty?

………

QUALITY CONTROL

16.   Do you have an internal quality control system?

☐ No   ☐ Yes

17. Do you apply the following quality controls?

Straight line controls[6]   ☐ No   ☐ Yes *Please specify the frequency* …………………………………..

Regression coefficient   ☐ No   ☐ Yes

Linearity coefficient   ☐ No   ☐ Yes

Slope   ☐ No   ☐ Yes

Blanks   ☐ No   ☐ Yes *Please specify the frequency* …………………………………..

Quality controls   ☐ No   ☐ Yes *Please specify the frequency* …………………………………..

Blind samples   ☐ No   ☐ Yes *Please specify the frequency* …………………………………..

Duplicate samples   ☐ No   ☐ Yes *Please specify the frequency* …………………………………..

18. Do you use certified reference materials?

☐ No   ☐ Yes *Please specify the manufacturer and concentration: ………………*

……………………………………………………………………………………………………….

19. Do you use reference materials?

☐ No   ☐ Yes *Please specify the manufacturer and concentration: ……………….*

……………………………………………………………………………………………………….

20. Do you use calibrated/verified equipment?

☐ No   ☐ Yes

21. Do you have an annual equipment calibration plan and programme?

☐ No   ☐ Yes

---

[6] This is related to the frequency of calibration to confirm that the curve parameters fulfil the validation criteria defined.

22. Do you have records of sample storage conditions, when necessary?

☐ No ☐ Yes *Please specify the parameter/s……………………………………*.

23. Do you have an annual intercomparison programme?

☐ No ☐ Yes

24. How often do you participate? (Please indicate organizer and number of times per year)

………

25. The evaluation of your interlaboratory results is based on:

z-score ☐ No ☐ Yes

E_n number ☐ No ☐ Yes

z'-score ☐ No ☐ Yes

Zeta-score ☐ No ☐ Yes

E_z-score ☐ No ☐ Yes

Other ☐ No ☐ Yes *Please specify ………………………………………………………………..*

Evaluation criteria

Question Number	Explanation
6	YES: Accredited laboratories must be considered expert laboratories; however, in this case, most of the other questions must be answered in the affirmative for this to be considered.
	NO: The proficiency of the laboratory can be estimated on the basis of the remaining questions.
7	YES: The first step when developing a particular validation should be the drafting of the general procedure for that validation.
	NO: If the method for the determination of mercury (question 9) has been validated, an appropriate proficiency could be aceptable.
8	YES: The second step when developing a particular validation should be the drafting of the particular procedure for mercury validation. This should be the initial step in the validation. NO: If the method for the determination of mercury (question 9) has been validated, an appropriate proficiency could be aceptable.
9	YES: A validated method is a necessary step when evaluating the proficiency of the laboratory. In addition, if quality control questions are appropriately answered, and question 10 offers suitable statistical parameters, the laboratory performance could be considered sufficient.
	NO: The laboratory should be able to validate the method. As a minimum, it is desirable to obtain statistical parameters for accuracy and the limit of quantification.
10	These values allow the performance of a laboratory to be evaluated. A comparison between different laboratories allows the reliability of each one to be determined.
11	This question allows the statistical proficiency of the laboratory to be determined.
12	This question allows the statistical proficiency of the laboratory to be determined.
13	This question allows the statistical proficiency of the laboratory to be determined.
14	This question allows the statistical proficiency of the laboratory to be determined.
15	This question allows the statistical proficiency of the laboratory to be determined. In this particular case, the possibility of underestimating the uncertainty in the measurement must be evaluated as this could affect the ability to obtain comparable results.
16	YES: It is necessary to evaluate the scope of internal quality controls in order to ensure that any deviation will be detected.
	NO: The first step to be able to trust in the reliability of the results must be to have an internal quality-control system.
17	YES: It is not necessary to implement all controls, but a higher number of controls ensures better results.
	NO: The laboratory should try to implement at least some of the controls, for example a calibration curve control and some type of sample control.
18	YES: The use of certified reference material ensures an assigned value. Possible manipulations (dilutions…) must be considered in order to obtain the real final value in every case.
	NO: As a minimum, reference materials (question 19) must be used.
19	YES: The laboratory can use suitable materials provided that these materials have been appropriately characterized. NO: This question can be ignored if question 18 is affirmative.

Question Number	Explanation
20	YES: Equipment calibration ensures instrumental repeatability and avoids equipment-related errors.
	NO: Calibration is the first step in any equipment control. No measurements should be performed prior to the calibration of critical equipment.
21	YES: An annual calibration plan and programme ensures that all equipment is operating correctly. Intermediate verifications should be carried out when necessary.
	NO: All equipment must be calibrated before the analysis is conducted.
22	YES: Measurement traceability is essential for suitable control of the environmental conditions.
	NO: Temperature, humidity and other aspects must be monitored when necessary. If not, the final results should be considered to be unreliable.
23	YES: Annual intercomparison programmes show the willingness of the laboratory and must be considered to be a favourable answer.
	NO: Only long-term participation provides the laboratory with an effective tool for evaluating its results.
24	A long participation time must be evaluated positively, irrespective of the results of this participation.
25	The z-score could be an insufficient means of determining the proficiency of the laboratory. Additional methods show a better capacity of the laboratory.

The criterion for evaluating laboratories should be based on the information collected from questions 6–25 of this questionnaire. However, this criterion can vary and can be applied more or less strictly depending on specific requirements and situations. In view of this, the following critera can be applied.

- Laboratories reporting negative answers to questions 16 or 20 should be automatically excluded.

- Laboratories reporting fewer than nine positive answers must improve their quality system. In particular, method validation should be the final goal for all participants, and it is strongly recommended to obtain quality control criteria from this validation.

- For laboratories reporting fewer than 18 positive answers, special attention should be paid to the answers to questions 10–15 and 24 as these allow proficiency to be evaluated and therefore the laboratory to be considered as a candidate.

- A positive response to more than 18 questions indicates a good analysis proficiency profile. The laboratory can, therefore, be evaluated as a candidate for performing the analysis. However, it should participate in the specific interlaboratory comparison exercises for the WHO study.

Annex 1 REGISTRATION OF RECEPTION OF A SAMPLE

ID code ☐

---

1. ORIGIN OF THE SAMPLE:
  Centre:
  City/Country:
  Date of sampling:

2. SAMPLE RECEIVED:

☐ Urine

☐ Hair

Signature of reviewer:

3. SAMPLE RECEPTION:

DATE (*dd/mm/yy*) ☐☐☐   TIME (*hh:mm*) ☐☐

  A) PACKAGING

☐ **NO PROBLEMS DETECTED**

☐ **PROBLEMS DETECTED:**

☐ Packaging damaged
☐ Cooling agents defrosted
☐ Others:_____

  B) SAMPLES

☐ **NO PROBLEMS DETECTED**

☐ **PROBLEMS DETECTED:**

☐ Spilled sample/broken vessel
☐ Insufficient amount/volume (*specify the matrix:* _____)
☐ Inconsistency in the ID codes
☐ Others:_____

  C) DOCUMENTS

☐ **NO PROBLEMS DETECTED**

☐ **PROBLEMS DETECTED:**

☐ Absence of the registry of collected samples
☐ Absence of the hair sampling questionnaire
☐ Absence of the urine sampling questionnaire
☐ Absence of the study questionnaire
☐ Inconsistency in the ID codes
☐ Others:_____

4. DATE OF STORAGE/BIOBANKING:

5. COMMENTS:

ID CODES FOR RELATED SAMPLES

URINE ☐   HAIR ☐

Annex 2 REGISTRY FOR COLLECTED SAMPLES

ORIGIN:	CENTRE: ADDRESS: CITY/COUNTRY: CONTACT (NAME AND PHONE):		
ID code	**Sampling date** *(dd/mm/yy)*	**Sampling questionnaire attached***(Y/N)*	**Comments**

Shipment date: Name and signature of fieldworker: