DESIGN AND OPERATION OF AN INTERLABORATORY COMPARISON SCHEME

R.M. VOICULESCU, M. C. OLTEANU, V.M. NISTOR

Institute for Nuclear Research Pitesti, Romania, ramona.voiculescu@nuclear.ro

ABSTRACT

The *competence* of laboratories is assessed by two complementary techniques. One of the techniques is the on-site evaluation following the requirements of *ISO/IEC 17025:2005.* The other one implies the *proficiency testing* which involves the determination of laboratory performance by means of *interlaboratory comparisons*, whereby the laboratory performs practical tests and their results are further compared with those of other laboratories.
The *paper treats* one of the most important topics of the proficiency testing – *the interlaboratory comparison (ILC).* There will be presented the need, the purpose and the main objectives of an ILC and also a typically situation where an *interlaboratory comparison exercise (for radio-analytical methods) was planned. A fully description of the design and operation of an ILC scheme* is the main purpose of this paper. A special attention will be given to the *data analysis and evaluation of interlaboratory comparison* scheme results.

Key words: interlaboratory comparison, testing scheme, item

Introduction

It is known that today is providing increasing importance to the study and assessment of the state, quality and health of the environment. For this reason many organizations are involved in such activities rely on the quality of the information provided and on the precision and accuracy of the data on which the information is based.
The results of analytical measurements play a vital role in our daily lives because analytical data may be the basis upon which economic, legal or environmental management decisions are made. Therefore, they are essential in international trade, environmental protection, safe transportation, law enforcement, consumer safety and the preservation of human health. [1]
It is important that such measurements are accurate, reliable, cost effective and defensible. An important role in this field is played by worldwide laboratories, involved in the production of environmental data which in many cases leading to wider assessments. Because of the need to base scientific conclusions on valid and internationally comparable data, it is indispensable to ensure the quality of the data produced by each laboratory. The need for ongoing confidence in laboratory performance is not only essential for

laboratories and their customers but also for other interested parties, such as regulators, laboratory accreditation bodies and other organizations that specify requirements for laboratories (**Figure 1**). So, there is a growing need for proficiency testing[1] – PT (comparative testing). Proficiency testing is an important way of meeting the requirements of ISO/IEC 17025:2005 [2] in the area of quality assurance of laboratory results.
The PT is a method for regularly assessing the accuracy of the analytical data produced by the laboratories of particular measurements. The laboratories conduct the test under routine conditions, and report the result to the organizer by a deadline. Proficiency testing by interlaboratory comparisons[2] (ILC) is used to determine the performance of individual laboratories for specific tests or measurements.

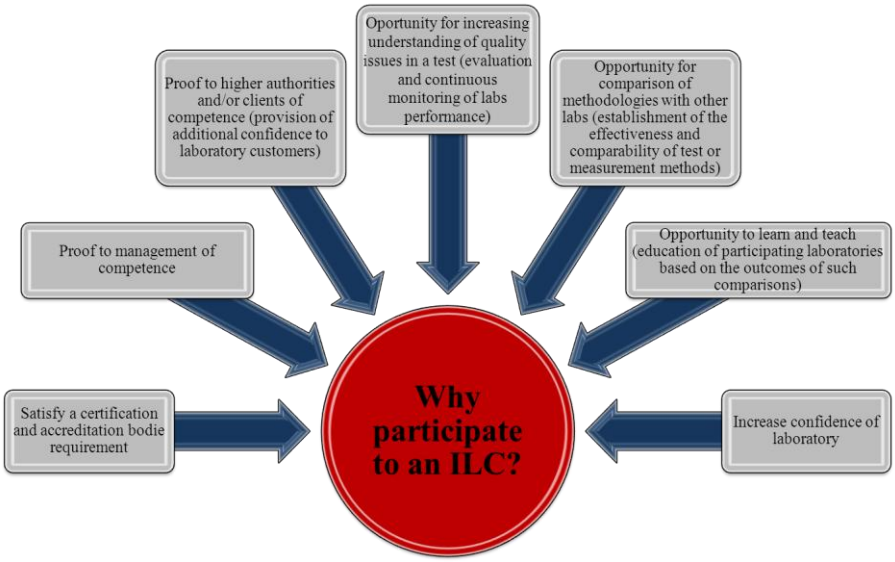

Figure 1 *Reasons for laboratories to participate to an ILC*

In this context, the paper will present the requirements of an interlaboratory comparison concerning the following three important topics of an ILC: testing scheme, participants and test items. A special attention will be given to the data analysis and evaluation of ILC scheme results. Also, a fully description of the design and operation of an ILC scheme will be provided by presentation of a typically example of ILC exercise (for radio-analytical methods).

1. Interlaboratory comparison – requirements of an ILC

Proficiency testing has become an essential aspect of laboratory practice in all areas of testing, calibration and inspection. PT schemes vary according to the needs of the sector in which they are used, the nature of the proficiency test items, the methods in use and the number of participants. However, in their simplest form, most proficiency testing schemes possess the common feature of comparison of results obtained by one laboratory with those obtained by one or more different laboratories. It is known that successful participation in an interlaboratory comparison is one of the necessary requirements of a laboratory

[1] ***Proficiency testing(PT)*** – evaluation of participant performance against pre-established criteria by means of interlaboratory comparisons; Interlaboratory comparison – organization, performance and evaluation of tests or measurements on the same or similar test items by two or more laboratories in accordance with predetermined conditions
[2] ***Interlaboratory comparison*** – organisation, performance and evaluation of test on the same or similar test items by two or more laboratories in accordance with predetermined conditions. Note – in some circumstances, one of the laboratories involved in the inter-comparison may be the laboratory, which provided the assigned value for the test item. [3]

accreditation or authorization covering at the same time the scope of the laboratory's accreditation in a useful and cost-effective manner. [4, 5, 6].
Although it was emphasized the importance of the PT there is no international organization, coalition, cooperation developed by and for proficiency testing providers, excepting EQALM[3] (which addresses to a narrow field) and ILAC[4]. ILAC created guidelines (Guides 13 and 43) for accreditations bodies, created the proficiency testing consultation group and worked with ISO to develop quality requirements for PT providers. Proficiency testing involves the use of interlaboratory comparisons for the determination of laboratory performance but not only. Interlaboratory comparisons are widely used for a large number of purposes and their use is increasing internationally. Typical purposes for interlaboratory comparisons are presented in **Figure 2**.

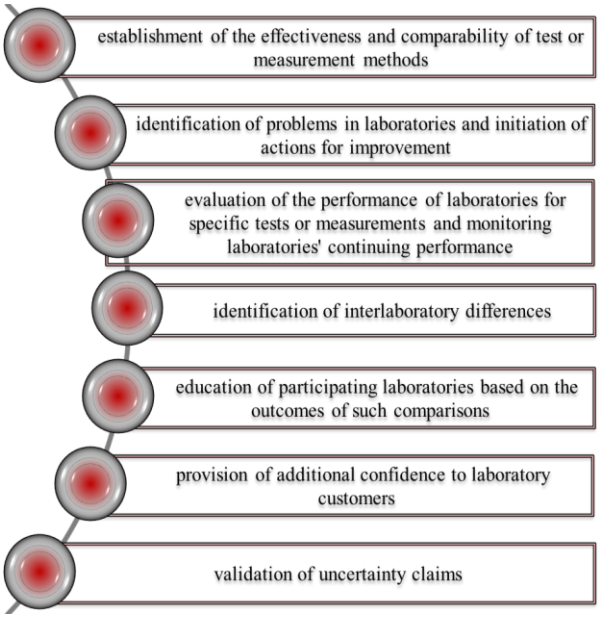

Figure 2 *Purposes for interlaboratory comparisons*

Irrespective of purpose, when an ILC it's organized, there must be considered at least the following elements:

a) Scope of the ILC;
b) Roles and responsibilities (establishment of the coordinator(s) and the test item provider(s), selection of participants);
c) Choice of ILC scheme (design and operation of testing scheme);
d) Test items (preparation, homogeneity and stability);
e) Statistical design;
f) Methods for data analysis and evaluation of ILC testing scheme results and Reports.

[3] EQALM – European Organisation for External Quality Assurance Providers in Laboratory Medicine; EQALM is an European group organisation involved in the external quality assessment of laboratory medicine services.
[4] ILAC – International Laboratory Accreditation Cooperation; is an international cooperation of laboratory and inspection accreditation bodies formed more than 30 years ago to help remove technical barriers to trade.

1.1. Testing scheme

The ILC scheme may be designed and operated in one or more rounds for a specified area of testing, measurement, calibration or inspection. This might cover a particular type of test, calibration, inspection or a number of tests, calibrations or inspections on test items. There are three basic types of laboratory examinations: *quantitative*, *qualitative and interpretive*. The results of a *quantitative measurement* are numerical and are reported on an interval or a ratio scale. Tests for quantitative measurement may vary in their precision, trueness, analytical sensitivity, and specificity. In quantitative proficiency testing schemes, numerical results are usually analyzed statistically. The results of *qualitative tests* are descriptive and reported on a categorical or ordinal scale, e.g. identity of micro-organisms, or by identification of the presence of a specific measurand (such as a drug or a grading of a characteristic). Assessment of performance by statistical analysis may not be appropriate for qualitative examinations. In *interpretive tests*, *the testing item* is a test result (e.g. a descriptive morphology statement), a set of data (e.g. to determine a calibration line) or other set of information (e.g. a case study), concerning an interpretative feature of the participant's competence.

Also the testing schemes may be:

- quantitative scheme – where the objective is to quantify one or more measurands of the proficiency test item;
- qualitative scheme – where the objective is to identify or describe one or more characteristics of the proficiency test item;
- sequential scheme – where one or more proficiency test items are distributed sequentially for testing or measurement and returned to the proficiency testing provider at intervals;
- simultaneous scheme – where proficiency test items are distributed for concurrent testing or measurement within a defined time period;
- single occasion exercise – where proficiency test items are provided on a single occasion;
- continuous scheme – where proficiency test items are provided at regular intervals;
- sampling – where samples are taken for subsequent analysis;

data transformation and interpretation – where sets of data or other information are furnished and the information is processed to provide an interpretation (or other outcome) [4].

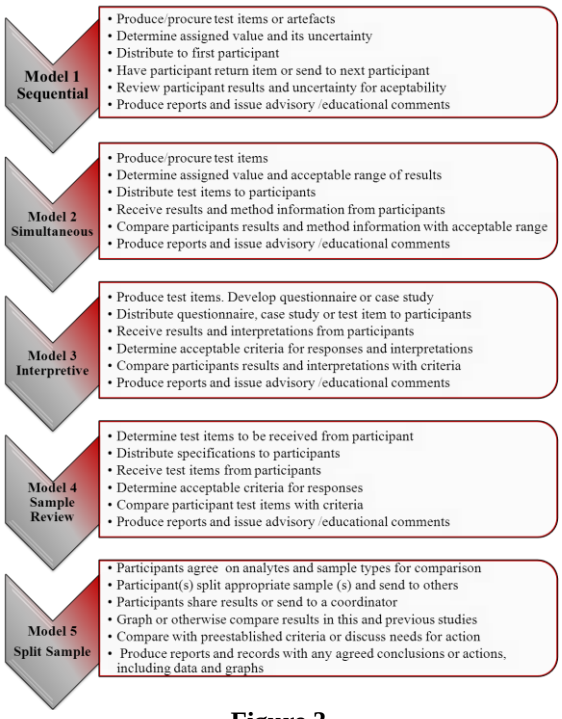

Figure 3
Examples of common types of testing schemes

The design of an ILC scheme must contain the following important elements:

- *Planning* – the coordinator will document a plan before commencement of the ILC testing scheme that addresses the objectives, purpose and basic design of the ILC testing scheme; this planning must contain a lot of information (the requirements concerning the planning are found in ISO 17043 – planning section), such as: identification data about the testing provider, requirements for participations, data concerning the test items and measurand, calendar etc.
- *Preparation of test items* – the coordinator and/or expert group will establish and ensure appropriate acquisition, collection, preparation (number of test items), handling, storage and, where required, disposal of all ILC test items;

- *Homogeneity and stability of test items* – the requirements in this sub-clause are intended to ensure that every participant receives comparable proficiency test items, and that these proficiency test items remain stable throughout the proficiency testing. In some cases, it is not feasible for proficiency test items to be subjected to homogeneity and stability testing. Such cases would include, for example, when limited material is available to prepare proficiency testing items;
- *Statistical design* – statistical designs shall be developed to meet the objectives of the scheme, based on the nature of the data (quantitative or qualitative, including ordinal and categorical), statistical assumptions, the nature of errors, and the expected number of results;
- *Assigned values* – the ILC coordinator shall document the procedure for determining the assigned values for the measurands or characteristics in a particular testing scheme. This procedure shall take into account the metrological traceability and measurement uncertainty required to demonstrate that the testing scheme is fit for its purpose [4,7].

1.2. Participants

Laboratories (and other types of participants) must participate to ILC testing schemes that are appropriate for their scope of testing. The ILC testing schemes selected should comply with the requirements of this International Standard (ISO/IEC/17045). Also, in order to participate to an ILC testing scheme, the participant should consider several aspects: the tests involved in the ILC match the types of tests, measurements or calibrations performed by the participant?, all details about the scheme design, procedures for establishment of assigned values, instructions, statistical treatment of data, and the final summary report are available? etc. [7, 8, 9] After all this are in place, the participant is ready to receive the test item(s) along with specific documentation (the most important being the instructions) and then the test can begin.

1.3. Test items

First of all, the test items must be prepared in accordance with the plan described in ISO/IEC/17043. It is advisable that the items provider give consideration to the preparation of sufficient numbers of samples.
The items provider is responsible with appropriate acquisition, collection, preparation, handling, packaging, transport and storage of samples. Moreover, where required, he must assure disposal of all ILC test items after the end of the tests. Another important aspect related to the test items is the demonstration of "sufficient homogeneity" (with valid statistical methods including a statistically random selection of a representative number of samples) and "stability" (stability is normally checked to ensure that the measurand(s) did not change during the course of the round). The entire algorithm for homogeneity and stability check is presented in Annex B of ISO 13528:2005 [10].

2. Assessment of the results

Assessment of results in interlaboratory comparisons, designed for purposes other than proficiency testing, has some features. This type of interlaboratory comparison can be planned and carried out among the laboratories themselves, or among the laboratories of one organisation. The advantage of this is the availability of results in a shorter time and the lower costs. Furthermore, they have the advantage that they can be applied to the specific problems of laboratories. A precondition for the recognition of interlaboratory comparisons is that the provider of the intercomparison (the coordinator) should clearly state in their programmes the assigned values according to ISO 13528:2005. Results received from participants are recorded and analysed by appropriate methods, established in the planning of ILC scheme. Data analysis shall generate summary statistics and performance statistics, and associated

information consistent with the statistical design of the ILC testing scheme. Methods of evaluation used must be valid (responsibility of coordinator) and meet the purpose of the ILC testing scheme. The methods shall be documented and include a description of the basis for the evaluation. ILC test results can appear in many forms, spanning a wide range of data types and underlying statistical distributions. The statistical methods used to analyse the results need to be appropriate for each situation. Some of the methods in ISO 13528, especially for homogeneity and stability testing, are modified slightly in the IUPAC Technical Report. The methods presented in the referenced documents (especially ISO 13528:2005) cover the fundamental steps common to nearly all ILC testing schemes:

- determination of the assigned value;
- calculation of performance statistics;
- evaluation of performance;
- preliminary determination of test item homogeneity and stability.

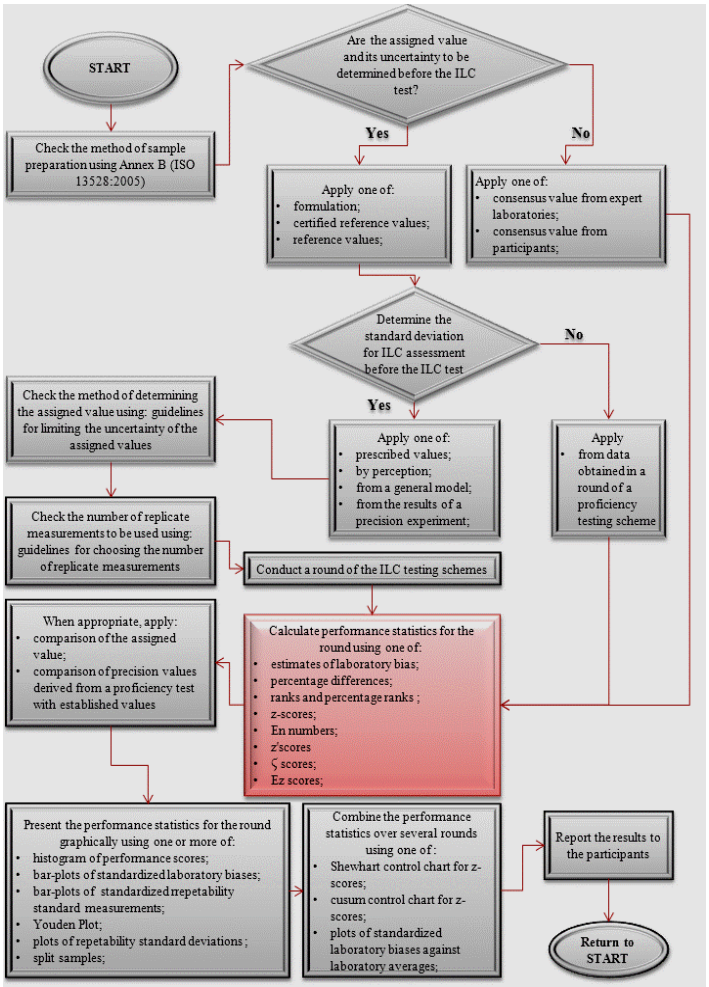

Figure 5 *Flowchart showing the activities requiring the use of statistical methods when operating a proficiency testing scheme. This flowchart is also applicable to an ILC*

2.1. Methods of evaluation and interpretation of results

For interlaboratory comparisons, which are organised or carried out by the laboratories themselves, an additional examination of the proper choice of the selected methods should be made by the assessment team. A precondition is that the laboratory organising the interlaboratory comparisons defines the assigned values. The methods for evaluation and interpretation of results are:

- Methods for performance statistics;
- Graphical methods for combining performance scores for several measurands from one round of an ILC test.

All these methods are listed in figure 5 and a large description of them is made in ISO 13528:2005.

2.2. Calculation of performance statistics

Before calculate the performance statistic (methods listed in the red square of fig. 5), all the steps described in the flow chart, fig. 6, must be followed. Calculating the performance statistics, it means:

- Estimate the laboratory bias;
- Calculate the percentage differences, the ranks and percentage ranks;
- Determine the *z* – score;
- Calculate *En* numbers, *z'* – scores, zeta-scores (ζ) and *Ez* – score.

ISO 13528:2005 give interpretation for all these parameters and more than that, there are presented calculation examples for each. The international standard shows some graphical methods for combining performance scores for several measurands from one round of an intercomparison test and also for several rounds of an ILC testing scheme.

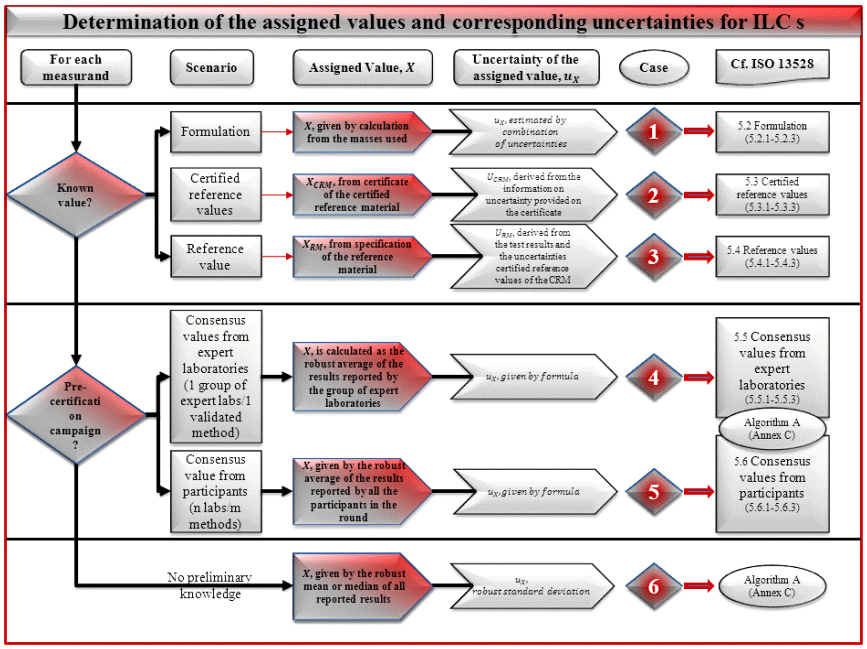

Figure 6 *ILC Strategy flow chart [11]*

3. ILC at INR Pitesti – A case study

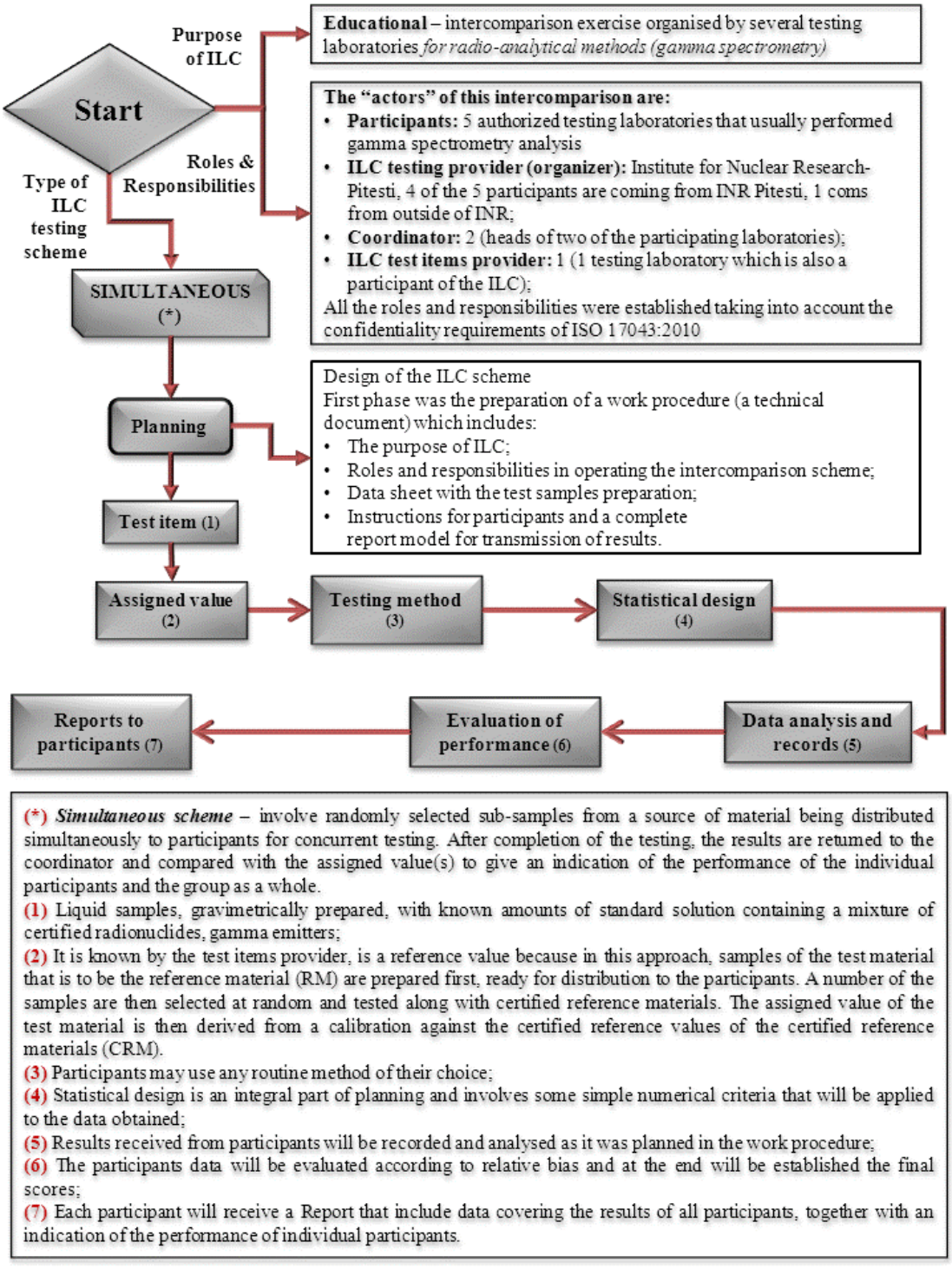

Figure 7 *ILC testing scheme for gamma spectrometry analysis*

Conclusions

In order to develop a successful PT plan for a testing laboratory, it is important to understand the documented PT requirements from the international standard generally and accreditation body particularly. The primary aim of proficiency testing is to provide a quality assurance tool for individual laboratories to enable them to compare their performance with similar laboratories, to take any necessary remedial action, and to facilitate improvement. Proficiency testing involves the use of interlaboratory comparisons for the determination of laboratory performance and it has so many benefits besides being a useful tool for accreditation. In the absence of an ILC provider, the design and operation of such scheme can be accomplished (trying to fulfill or adapt as many as possible the requirements of ISO/IEC 17043:2010) and is a real challenge for testing laboratories.
At the INR Pitesti, an interlaboratory comparison scheme was design by several testing laboratories, in order to establish the effectiveness and comparability of test and measurement method and also for education of participating laboratories based on the outcomes of such comparison. The ILC scheme is in progress and all the steps in her design were made taking into account the general requirements for PT from ISO/IEC 17043:2010, and in order to determine the performance of individual laboratories (for radio-analytical specific tests), it was developed a statistical model following the ISO 13528:2005 requirements.
All the participating labs are authorized by National Commission for Nuclear Activities Control, so the ILC wasn't done for accreditation but for demonstrating competence based on practical evaluation and for increase confidence of laboratory (feel good!).

References

[1] IAEA Analytical Quality in Nuclear Applications No. IAEA/AQ/6, *Proficiency Test of Ukrainian Laboratories Determination of Gamma Emitting Radionuclides*, VIENNA, 2009, p.1
[2] ISO/IEC 17025:2005 – General Requirements for the Competence of Testing and Calibration Laboratories
[3] ISO/IEC Guide 43:1997, Part 1: Development and Operation of Proficiency Testing Schemes (2nd Edition, 1997)
[4] ISO/IEC/17043:2010, Conformity assessment – General requirements for proficiency testing (first edition, 2010-02-01)
[5] BLQS, DMSc N 07.15.003, Policy and requirements for proficiency testing, interlaboratory comparison/laboratory's performance assessment in test, September 2009, revision 4
[6] Vladimir Kocourek, Interlaboratory Studies, Institute of Chemical Technology, Department of Food Chemistry and Analysis, Prague, 2012
[7] APLAC PT 002, Testing Interlaboratory Comparisons, March, 2008
[8] IFM, Quality Services. Tutorial for participant laboratories, 2004
[9] Michael A Noble, Chair, Clinical Microbiology Proficiency Testing Program, Department of Pathology and Laboratory Medicine, University of British Columbia Canada, International Standards for Proficiency Testing Providers
[10] ISO 13528:2005 – Statistical methods for use in proficiency testing by interlaboratory comparisons
[11] M. Thompson, S.L.R. Ellison, R. Wood – The International Harmonized Protocol for the proficiency testing of analytical chemistry laboratories (IUPAC Technical Report), in *Pure and Applied Chemistry*, Vol. 78, No. 1, pp. 145-196, 2006