Asia-Pacific
Economic Cooperation

Committee on Trade and Investment
Sub-Committee on Standards and Conformance

Asia-Pacific Legal Metrology Forum

APEC

Handbook on Traceability in Legal Metrology

March 2004

Asia-Pacific
Economic Cooperation

Committee on Trade and Investment
Sub-Committee on Standards and Conformance

Asia-Pacific Legal Metrology Forum

Handbook on Traceability in Legal Metrology

March 2004

APEC Project CTI 25/2003T

APLMF Symposium on Traceability in Legal Metrology

Prepared by:

Secretariat
Asia-Pacific Legal Metrology Forum
NMIJ/AIST Tsukuba Central 3-9
1-1-1 Umezono, Tsukuba, Ibaraki 305-8563, Japan
Tel: (81) 29 861 4362 Fax: (81) 29 861 4393
Email: e.sec@aplmf.org
Website: www.aplmf.org

For:

APEC Secretariat
35 Heng Mui Keng Terrace Singapore 119616
Tel: (65) 6775 6012 Fax: (65) 6775 6013
Email: info@mail.apecsec.org.sg
Website: www.apec.org

APEC#204-CT-01.4 ISBN 4-9901968-0-5

Contents

Foreword

Akira Ooiwa

APLMF President

National Metrology Institute of Japan

National Institute of Advanced Industrial Science and Technology

AIST Tsukuba Central 3-1, 1-1-1 Umezono, Tsukuba, Ibaraki 305-8563, Japan

This booklet is one of outcomes of the symposium with a title 'Traceability in Legal Metrology' that was held in 30th to 31st of October 2003 in Kyoto, Japan organized by Asia-Pacific Legal Metrology Forum (APLMF) with a fund supported by Asia-Pacific Economic Cooperation (APEC). On this occasion, I would like to extend my deepest thanks to all the participants and contributors from APEC member economies and from international and regional bodies especially to the secretariats of APEC and APLMF for their voluntary supports.

Main target of this symposium so called 'Traceability in Legal Metrology' was to design new desirable linkages between legislative demands for measurement and scientific metrological resources, aiming at establishing good partnerships among legal measurements and scientific standards. Good partnerships will realize 'Measurement Traceability in Legal Metrology' and consequently will lead to facilitate possible future mutual recognition in the area of international trade and relating legislative systems. In view of establishing desirable linkages, this symposium should be recognized as the important first step for the players from both of the areas, legal metrology and scientific standards, and through this new relationships great benefits will be provided in the fields of trade facilitations.

The focused subjects were as follows.

a) SI Traceability for Legal Metrology
b) Uncertainty analysis and confidence level of legal metrology
c) Accreditation in Legal Metrology
d) Difficult metrology fields to get traceability, such as analytical metrology, blood alcohol, toy safety, mass, and so on.

With the progress of international market integration relating with APEC/WTO activities, there have been expanding new technical fields that should be covered by public control. For example, a reliable measurement of undesirable chemical substances in foods has appeared one of significant issues for the safety of people's quality of life. Measuring methods in such serious objects should be authorized based on suitable internationally accepted technical regulations with common consent. In most cases of such newly demanded cases, the linkages among technical specialists and specific legal administrators are not satisfactory. This situation means that Legal Metrology needs technical partnerships with scientific standards by means of the idea 'Measurement Traceability' as the basis of measurement confidence.

In Asia-Pacific Region, we already have acknowledged activities in the scientific metrology, Asia-Pacific Metrology Program (APMP), and in the laboratory accreditation, Asia-pacific Laboratory Accreditation Cooperation (APLAC), those should be linked with APLMF demands. In view of these situations, this symposium was organized by APLMF with a support from APEC in order to collaborate with other specialist regional bodies, APMP and APLAC, in order to prepare a sure basis of metrological confidence in Legal Metrology.

I am really pleased to have this outcome from the first symposium of 'Traceability in Legal Metrology' and again deeply appreciate invaluable voluntary efforts of APEC and APLMF secretariats, especially their enthusiasm and patience even at the occasion when the difficult problems happened

February 23, 2004

Dr. Akira Ooiwa

APLMF President

Opening Address

John Birch AM

Past President of APLMF, Australia

I would like to congratulate Dr Ooiwa on his initiative in organising this Symposium on Traceability in Legal Metrology. The focus of the Symposium is both important and timely and it is a particularly appropriate topic to celebrate the centenary of the National Measurement Institute of Japan.

I would also like to thank him for organising this Symposium in Kyoto, the historical and cultural centre of Japan. It is always a joy to visit this city. Any study of the history and culture of a nation will provide insights into the development of metrology. This was brought home to me earlier this week when I attended an exhibition in the Nara National Museum of the Shoso-in Treasures. Of the sixty-six treasures from the eight century on display three related to metrology.

A wooden ruler 44.5 cm long with precise decimal graduations highlighted the role of metrology in trade and manufacture. A 766CE map of farmland highlighted the role of surveying in establishing land ownership and a ceremonial red stained ivory ruler which was offered to the palace emphasised the role of the State in setting the rules of the measurement system.

Traceability has been the defining principle of metrology for over 5000 years. The State found that it needed consistency of measurements to allow for their aggregation to provide the information needed to organise plan, defend and tax with efficiency. This consistency was achieved by requiring measurements to be derived from a single State standard.

Whilst traceability has served metrology well, changing circumstances requires some reinterpretation, in particular;

- Artefact standards have, with the exception of mass, been replaced by physical standards.
- Physical measurements have expanded to a wide range of phenomena and are used for many regulatory purposes, requiring traceability to an increasing range of physical quantities.
- Globalisation requires measurements to be accepted worldwide with the same degree of confidence as exists nationally.

These changes can result in increased disputation unless trust and confidence is established in the expanded measurement system. Legal metrology has an important co-ordinating role in establishing this trust and confidence.

A key responsibility of legal metrology is a legislative definition of traceability. This is currently being considered in the review of the OIML Document 1 Law on Metrology. A legislative definition of traceability provides;

- Legal standing for the national primary standards.
- A sound evidential basis for measurements.
- Defines the commitment of the government to the measurement system.

All of the issues I have mentioned will be discussed over the next two days so I am sure we can look forward to an interesting and productive symposium.

Accreditation and Traceability in Legal Metrology

SETA Katuo

Chief Executive, International Accreditation Japan (IAJapan)

National Institute of Technology and Evaluation (NITE)

2-49-10 Nishihara, Shibuya-ku, Tokyo151-0066, Japan

Accreditation is defined as assessment of "conformity assessment bodies (CAB)", which includes testing and calibration laboratories and also inspection and certification bodies. In many of international standards and guidelines applied to accreditation, traceability is required to confirm technical competence of measurement or to obtain a measurement result, all over the world. Therefore, we can utilize the accreditation to demonstrate the traceability and reliability of measurements in various aspects, such as quality control of products, monitoring environments and so on.

The accreditation bodies, ABs, are trying to expand the aspects where the accreditation works practically. Legal metrology is one of the important targets, since it requires confirming traceability and technical competence of laboratories and manufacturers of measuring instruments.

1. Introduction

It has been recognized that measurement traceability is one of the key issues in various fields of measurements including legal metrology. Recently not only to establish traceability but also to demonstrate traceability is required especially to reduce and/or to remove technical barrier of trade, TBT. More general, conformity assessment has been activated for these decades since it has key role to reduce TBT. Accreditation is thought to be a practical tool to endorse reliability of conformity assessments including the measurement traceability.

Here, the current view of conformity assessment and accreditation is introduced in section 2 and document standards and guidelines for them are listed in section 3. Relationship between traceability and accreditation is discussed in section 4 and international situation of accreditation is shown in section 5. Also, the possibility of accreditation to be utilized in legal metrology is suggested in section 6.

2. Current view of Conformity Assessment and Accreditation

The conformity assessment includes testing, calibration, inspection, certification and so on. These activities assess their objects directly;

- testing: determination of one or more characteristics of a given product, process or service,
- calibration: comparison between reference standards and measuring instruments or artifacts,
- inspection: judgment accompanied as appropriateness by measurement, testing or gauging,
- certification: procedure by which a third party gives written assurance that a product, process, or service conforms to specified requirements.

For a period, accreditation had been counted as one of the conformity assessment activities, however, it has been defined as "assessment of other conformity assessment activities" at

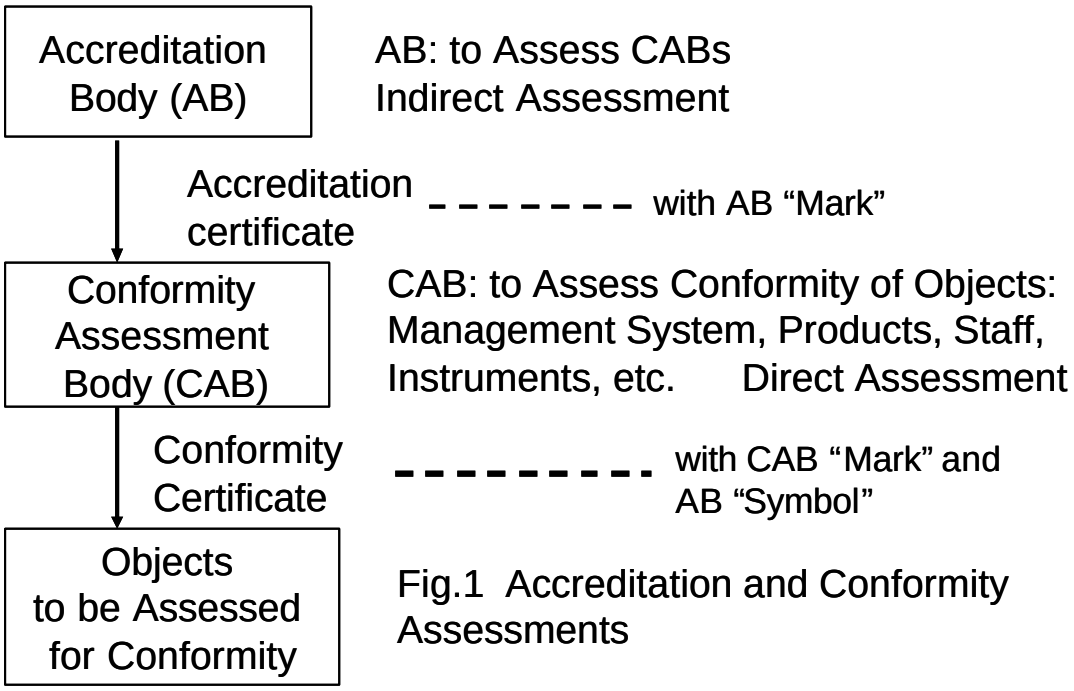


Fig.1 Accreditation and Conformity Assessments

present as shown in Fig.1. Accreditation Bodies (ABs) assess and accredit Conformity Assessment Bodies (CABs) and CABs demonstrate reliability of their activities through accreditation. Accreditation is "final and the highest stage of Assessment", and defined as below in international standards:

- Accreditation
  - procedure by which an authoritative body gives formal recognition that a body or person is competent to carry out specific tasks (ISO/IEC Guide 2: 1996)
  - third-party attestation that a *conformity assessment body* fulfils specified requirements and is competent to carry out specific conformity assessment tasks (ISO/IEC 17011DIS)

- NOTE 1 Formal recognition of the conformity assessment body's competence is provided by an accreditation body

- Accreditation Body (AB)
  - authoritative body that performs accreditation

(ISO/IEC 17011DIS) DIS=Draft International Standard

while conformity assessments are defined in the same standards as below:

- Conformity Assessment
  - any activity concerned with determining directly or indirectly that relevant requirements are fulfilled (ISO/IEC Guide 2: 1996)
  - NOTE Typical examples of conformity assessment activities are sampling, testing and inspection; evaluation, verification and assurance of conformity (supplier's declaration, certification); registration, accreditation and approval as well as their combinations.
- Conformity Assessment Body (CAB)
  - body that performs conformity assessment services and that can be the object of accreditation (ISO/IEC 17011DIS)
  - For the purposes of this International Standard, CABs are organisations providing the following conformity assessment services: calibration, testing, inspection, management system certification, personnel certification and product certification.

Because of this characteristic, "one national accreditation body" has been developed in many countries/economies like National Metrology Institute (NMI) in metrology, while in some countries/economies plural ABs have been developed in various fields where accreditation is required.

3. Features and merits of laboratory accreditation

We can count features and merits of laboratory accreditation as below.

1) International Standards
   - Testing and Calibration Laboratories: ISO/IEC 17025
   - Accreditation Bodies: ISO/IEC Guide 58 (ISO/IEC 17011 in future)
     - Eligible and trained assessor with technical background
     - Transparent and objective assessment and decision making
     - Management system considering impartiality
     - Confidentiality
     - Application of Proficiency Test (PT)

2) Technical Conformity
   - Emphasis of Technical Competence (Cf: certification of ISO 9000)
     - ***Traceability***
       Requirements of traceability (unbroken chain of measurement values with information of uncertainties) to SI / NMI
     - Uncertainty
       Estimated quantitative reliability of measurement values
     - Method Validation
       Requirement to confirm if the measurement method gives correct value
     - Others (Personnel, Sampling etc.)
     - In order to secure conformity, it is required to notice reliability of measurement values with its uncertainties
3) Continuity
   - Assessment with international standards at initial accreditation
   - Periodical severance and participation in PT to maintain technical competence continuously after initial accreditation
   - Continuous reliability for results of conformity assessment as a result
4) Mutual Recognition Arrangement
   - Peer evaluation among ABs to confirm if ABs and CABs accredited satisfy requirements for conformity
   - Signature to the MRA as result
   - Duty of MRA members: to recognize equivalence of capability and results of other member ABs, to negotiate the government to accept accreditation of foreign member ABs
   - Follow up with periodical re-evaluation for MRA members
5) Indirect Results
   - Reduction of government task by using expert organization
     - Technical requirements in regulations and in purchase
     - Criteria for authorization, approval and license
   - Progress in reliability of laboratories by 3rd party assessment
     - Reduction of task for advertisement and exploitation
   - Reduction of user ' s tasks to accept certificates
     - Reduction in research and evaluation
   - Acceleration of trade
     - International reliability

4. Standards of Conformity Assessment and Accreditation

Conformity assessment requires its documented standards for assessment activities. Considering its purpose, these standards need consensus of world wide and ISO-CASCO (International organization for Standardization - Committee on Conformity Assessment) has worked to obtain the international consensus in this field, collaborating with IEC (International Electrotechnical Commission). The standards and guidelines for conformity assessment activities are listed in Table 1, where those for accreditation of CABs are also listed. These have been issued and amended by ISO-CASCO. Testing, calibration, inspection, certification of QMS (Quality Management System), EMS (Environmental Management System), products certification and staff certification are recognized as "Conformity Assessment", and the standards for accreditation of these activities are under process of unification to ISO/IEC 17011, which is planed to be fixed soon. On the other hand, proficiency test (PT) and reference material (RM) production are not include in conformity assessment, to be accredited by ABs operating ISO/IEC 17011 system.

Table 1 Conformity assessments, accreditation and their standards

Activities		Conformity assessment bodies (CAB)	Standards for assessment	Standards or Guidelines for CABs	Standards or Guidelines for ABs
Test *(ILAC-MRA)*		Testing laboratory	Standards for tests	ISO/IEC 17025	ISO/IEC G58 to 17011
Calibration *(ILAC-MRA)*		Calibration laboratory	Metrology Std. traceable to SI	ISO/IEC 17025	ISO/IEC G58 to 17011
Inspection		Inspection body	Standards for safety etc.	ISO/IEC 17020	ISO/IEC 17010 to 17011
Certifi-cation	QMS *(IAF-MRA)*	Certification body	ISO/IEC 9001	ISO/IEC G62 to 17024	ISO/IEC G61 to 17011
	EMS *(IAF-MRA)*		ISO/IEC 14000	ISO/IEC G66 to 17024	to ISO/IEC 17011
	Products		Standards for Products	ISO/IEC G65	ISO/IEC G61 to 17011
	Staff		Qualification of Staff	to ISO/IEC 17024	to ISO/IEC 17011
Proficiency Test		PT provider	Reference value	ISO/IEC G43-1, 2	One part of accreditation process ?
Reference Materials		RM producer	Specification of RM	ISO G34 or ISO/IEC 17025	Certification or Accreditation ?

5. Traceability utilizing Accreditation

The definition of "Traceability" is given in VIM (International Vocabulary of basic and general terms in Metrology: 1993), as "the property of a measurement or the value of a standard whereby it can be related to stated references, usually national or international standards, through an unbroken chain of comparisons all having stated uncertainties". Considering the meaning of "unbroken chain", it should be noted that traceability only exists when scientific rigorous evidence is collected on a continuing basis showing that the measurement produces documented results for which the total measurement uncertainty is quantified. In other words, it is required for "chain laboratories" to demonstrate maintenance of technical competence including estimation of uncertainties. For this purpose, laboratory accreditation applying to ISO/IEC 17025 is useful way.

As an example, Japanese traceability system including accredited laboratories is shown in Fig.2. In this figure, vertical solid lines from NMIs to Society show traceability through calibration and testing laboratories. "ASNITE", "JCSS" and "JNLA" are programs operated by IAJapan, broken lines with arrow show accreditation. The accredited laboratories are required to satisfy ISO/IEC 17025 and IAJapan is required to satisfy Guide 58 (ISO/IEC 17011 in near future) in operation of these accreditation programs. Because of the accreditation, the laboratories can demonstrate their technical competence easily, just showing accreditation certificates. The National Metrology Institutes (NMIs) are participating to their own global Mutual Recognition of Arrangement (CIPM-MRA), while the accreditation with "ASNITE-NMI" program is utilized to demonstrate quality system in their calibration services for participation in the CIPM-MRA. On the other hand, the NMIs are providing reference values in the proficiency testings (PT) of calibration laboratories, then, the PTs give important information in accreditation process.

IAJapan is an active member of International Laboratory Accreditation Cooperation (ILAC) and participates in its Mutual Recognition Arrangement (MRA) as shown in Fig.2. The international relation is shown in Fig.3. Not only NMIs, but also ABs should participate in the MRA to demonstrate national traceability system internationally as shown in Fig.3. There are many laboratories even in an economy and it is difficult to evaluate technical competence of a laboratory in a foreign economy when you received a calibration certificate or a test report issued by the laboratory. In the system shown in Fig.3, the reliability and technical competence of the laboratory are endorsed by NMIs and ABs, then, the NMIs and ABs are evaluated internationally in order to participate to their MRAs. If the reliability of these two MRAs, of CIPM and of ILAC, is enough high, it is easy to accept calibration certificates and test reports issued by accredited laboratories in other economies.

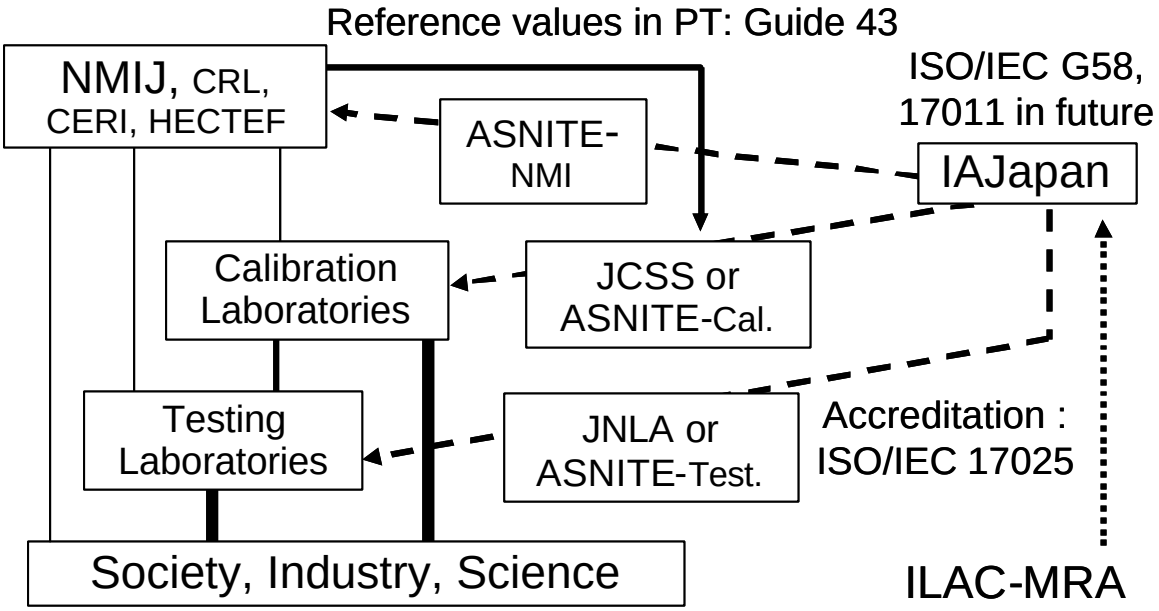


Fig.2 Traceability and Laboratory Accreditation System in Japan

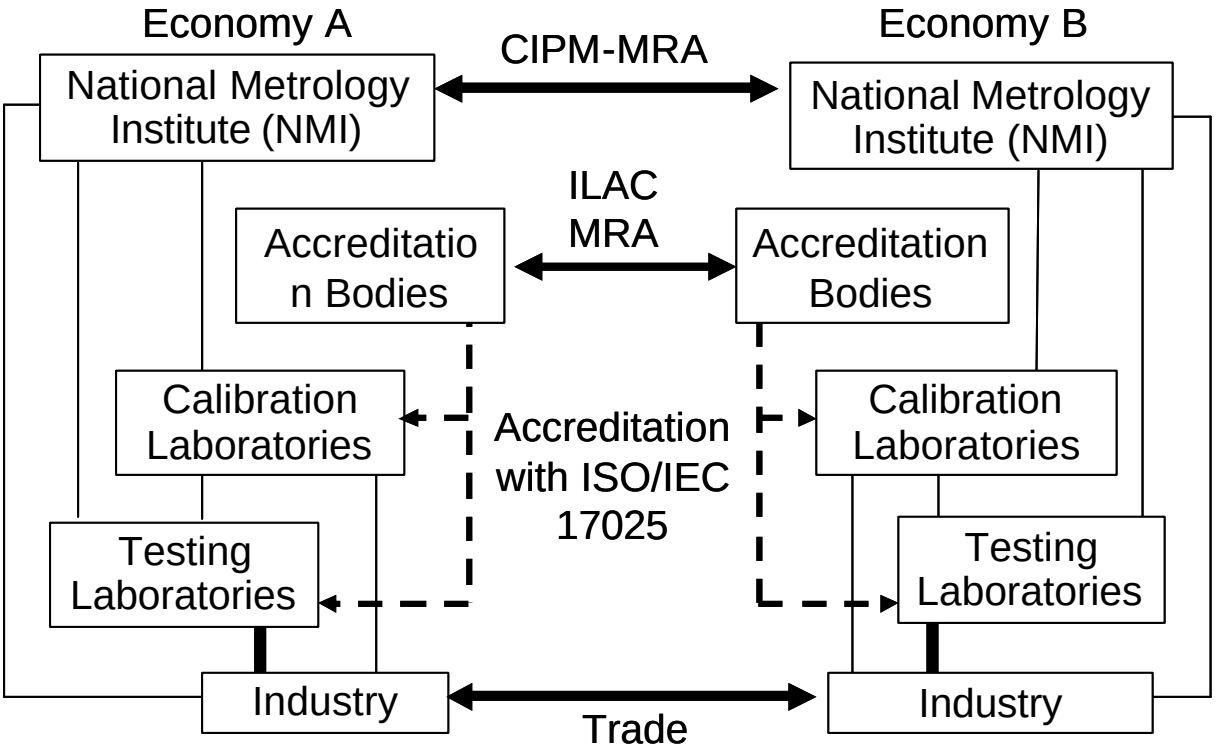


Fig.3 International Traceability System

6. ILAC and its MRA

ILAC is an international cooperation of laboratory accreditation schemes all over the world. It was founded about 20 years ago and was formalized as cooperation in 1996 when 44 national bodies signed a Memorandum of Understanding (MoU). The ILAC defined its role:

"ILAC is an international cooperation between the various laboratory accreditation schemes operated throughout the world. Founded twenty years ago, ILAC was formalized as a cooperation in 1996 when 44 national bodies signed a Memorandum of Understanding (MOU) in Amsterdam. This MOU provides the basis for the further development of the Cooperation and the eventual establishment of a multilateral recognition agreement between ILAC member bodies. Such an agreement will further enhance and facilitate the international acceptance of test data, and the elimination of technical barriers to trade as recommended and to support the World Trade Organization (WTO) TBT agreement."

The main purpose of ILAC is to establish a multilateral recognition agreement between ILAC member bodies. Such an agreement will further enhance and facilitate the international acceptance of test data, and the elimination of technical barriers to trade and support the World Trade Organization (WTO) TBT agreement. On 20 January 2003, ILAC was successfully incorporated and became an Association under the Dutch Law.

ILAC has arranged the Mutual Recognition Arrangement as mentioned in the previous section since 1996. In order to establish the MRA, ILAC has developed criteria and guidelines to participate the MRA. The accreditation body is required to operate quality system according with ISO/IEC Guide 58 to assess laboratories based on ISO/IEC 17025.

Each AB that participates in the ILAC MRA shall:

- maintain conformance with ISO/IEC G58, related ILAC guidance documents and, a few, but important supplementary requirements, and
- ensure that all accredited laboratories comply with ISO/IEC 17025 and related ILAC guidance documents.

When an AB applies to participate to the MRA, the AB is to be evaluated by member bodies, to confirm if the AB satisfies ILAC requirements. In the evaluation process, 3 - 6 evaluators from member bodies review quality documents of the accreditation body to be reviewed, then, visit the body to evaluate their accreditation activities for about a week.

The steps of evaluation process for MRA applicant are:

1. Evaluation team: Evaluators of senior staffs
2. Review of Quality documents of the applicant
3. Evaluation visit:
    - Applicant AB itself for documents, records and staffs
    - Witness for assessors, accreditation process
    - Accredited laboratories and NMIs
4. Review report to MRA council
5. Decision making at MRA council

Usually, the evaluation is not organized by ILAC itself but by regional accreditation cooperation organizations, Asia Pacific Laboratory Accreditation Cooperation (APLAC), European cooperation for Accreditation (EA) and South African Developing Cooperation of Accreditation (SADCA). If there is not suitable regional organization for an AB, ILAC organize the evaluation process by itself. Now, 44 ABs of 35 countries/economies participate in the ILAC MRA. More details are shown in the ILAC website: http://www.ilac.org/.

For laboratory accreditation, ILAC MRA has been established, while IAF (International Accreditation Forum) has established their MultiLateral recognition Agreement (MLA) for certification bodies of quality management as shown in Table1. However, the whole CAB activities have not yet been covered by these two MRAs. Now, the MRA for accreditation of inspection bodies is under discussion in both international organizations, while several international organizations, including ILAC and IAF are collaborating to obtain consensus for certification and/or accreditation of reference material producers. In the table 2, the scope of MRAs is shown for global and regional organizations of accreditation bodies.

Table2 Scope of MRAs among accreditation bodies in global or regional level

	Testing & Calibration	Inspection	Certification
Global	ILAC	Both ?	IAF
Asia-Pacific	APLAC	APLAC	PAC
Europe	EA	EA	EA
America	IAAC	IAAC	PAC
South Africa	SADCA	SADCA	SADCA

The membership of ILAC includes MRA members, MRA candidates, regional bodies, and so on, as below:

- 44 Full Members (Signatories to the ILAC MRA) representing 35 economies;
- 15 Associates representing 15 economies;
- 19 Affiliates representing 16 economies;
- 4 Regional Cooperation Bodies;
  - APLAC (Asia Pacific Laboratory Accreditation Cooperation)
  - EA (European Cooperation for Accreditation)
  - IAAC (InterAmerican Accreditation Cooperation)
  - SADCA (South African Developing Cooperation of Accreditation)
- 1 National Coordination Body; DAR (DE)
- 19 Stakeholders. (JLA etc.)

At the last ILAC-GA in Bratislava, it was decided to commence "ILAC-MRA" symbol scheme, where ABs of ILAC-MRA members may issue "Accreditation Certificates" with ILAC-MRA symbol and laboratories, accredited by an AB of ILAC-MRA members, may issue conformity assessment certificates with the accreditation symbol and ILAC-MRA symbol. The ILAC-MRA symbol could assist users to recognize the reliability of the conformity assessment certificates.

7. Other international activities of ILAC

The ILAC has developed various collaborations with other international organizations and has signed MoUs as below:

- United Nations Industrial Development Organization (UNIDO) / International Organization for Standardization (ISO) – 2000/2001
- Comité International des Poids et Mesures (CIPM) - 2001
- Industry Cooperation on Standards and Conformity Assessment (ICSCA) – 2002

ILAC and IAF are currently working with ISO to develop a new tripartite MoU. This process was further facilitated by the presence of the ISO/CASCO Secretary at the recent Bratislava meetings. Following a review of the draft MoU document by the executive bodies from all three organisations and some further development work in Bratislava a new draft is now being finalised.

At present, ILAC arranges formal liaisons for the organizations listed below:

- *ISO* – Plenaries, Chairman's Advisory Group, Joint Working Group (IAF/ILAC), and various other CASCO Working Groups, Committee on Reference Materials (REMCO) and various Technical Committees
- *IEC* (International Electrotechnical Commission): Conformity Assessment Board
- *CIPM* (Comité International des Poids et Mesures): Bureau International des Poids et Mesures (BIPM), associated Committees like the Joint Committee on Traceability in Laboratory Medicine (JCTLM) and the Joint Committee of the Regional Metrology Organisations and the BIPM (JCRB)
- *OIML* (International Organization of Legal Metrology)
- *ICSCA* (Industry Cooperation on Standards and Conformity Assessment)
- *UNIDO* (United Nations Industrial Development Organization)
- *NCSLI* (National Conference of Standards Laboratories International)
- *EASC* (Euro-Asia Council for Standardization, Metrology and Certification)
- *WTO* (World Trade Organization)

8. Domestic situation in each economy

Because of historical, economical, political or cultural reason, surrounding situation of accreditation bodies strongly depends on each economy. For the number of ABs in a economy, many of European economies has one national AB, while a number of accreditation bodies in USA, Japan and Germany. The outlines of situation are:

- EU Manifestation
  - technical final level, one system in an economy, non-competition, non-commercial
  - Under discussion in non-competition policy divisions
- One system in an economy
  - One system with plural ABs:
    - Italy: SINCERT (certification), SINAL (testing), SIT (calibration)
    - Belgium: BELCERT (certification), BERTEST (testing), BKO/OBE (calibration)
    - Australia: JAS-ANZ (certification), NATA (testing and calibration)
- Competition between ABs
  - USA (testing/calibration: NVLAP, A2LA, ICBO etc., 12 ABs in NACLA; certification: ANSI-RAB joint program)
    - Germany (DKD, DAP etc., about 20 ABs in DAR)
    - Japan (calibration: IAJapan, JAB; testing: IAJapan, JAB, JCLA, VLAC; certification: JAB, JASC, planned by IAJapan)

Also, the legal status depends on each economy:

- Central government sector
  - HKAS (HK), NVLAP (US), BELCERT• BELTEST• BKO/OBE (BE), BMwA (AT)
  - JASC (JP)
- Government Agency
  - IANZ (NZ), KOLAS (KR), NA (NO), DSM (MY), INMETRO (BR), ISRLAC (IL), SAS (CH), FINAS (FI), SWEDAC (SE), LA (LV), PCA (PL), DANAK (DK)
  - IAJapan/NITE (JP)
- Private Sector
  - UKAS (UK), COFRAC (FR), RvA (NL), NATA (AU), A2LA (US), EAK (EE), RENAR (RO), SINAL (IT), TURKAK (TR), SA (SI), CAI (CZ)
  - JAB, JCLA, VLAC (JP).

Many of ABs of private sectors, however, have specified relation with the governments as below:

- UKAS (UK): MoU with Ministry of Industry and Trade
- COFRAC (FR): considering contract and/or MoU

- RvA (NL): Contract with the government
- NATA (AU): Recognized by the government as the unique AB
- A2LA (US): NA
- EAK (EE): Contract with the government
- RENAR (RO): Recognized with (economy and trade) ministerial ordinance
- SINAL (IT): Recognized with (industry, trade and transportation) ministerial ordinance
- TURKAK (TR) : Founded and invested for accreditation by law
- SA (SI) : Invested for accreditation by law
- CAI (CZ): Commission for accreditation by the government.

9. Acceptance of accreditation and its MRA by regulators

The most important user of accreditation is regulators. The ILAC surveyed for the acceptance of accreditation and its MRA in 2002. The results show positive situation as total, all reply said "yes" for acceptance by regulators as below:

- Situation of acceptance of LA by regulatory bodies
  - LA is accepted and / or required legally in regulations
  - MRA members are accepted or not
- Replied by 17 countries/economies
  - All countries/economies accept LA but details depend on each country/economy
  - LA is required in 16 countries/economies
  - MRA members are accepted in 13 countries/economies.

The "good examples" are shown below:

- Singapore: SAC-Singlas
  - LA is accepted in " Registration system to protect consumers " : requirements on test reports: issued by laboratories which are accredited by SAC-Singlas or by its MRA partners (JAB and JNLA are listed as examples), to satisfy guide 25
- France: COFLAC
  - In mandatory regulations, LA is accepted and/or required for inspection of water, wine, test of atmosphere, BSE, asbestos etc.
- Finland: FINAS
  - In 30 mandatory regulations, at least, LA is referred (accepted or required) (Ministry of Trade and Industry 19, Ministry of Transportation 3 etc.)
  - In mandatory regulations:
    - Food testing: accredited by FINAS or other ABs which satisfy guide 58
    - Animal testing: judged with EN45001 or guide 25 by FINAS or other ABs which satisfy guide 58

- Water testing: to satisfy ISO/IEC 17025 and accredited by an AB, for example, FINAS

- Denmark: DANAK
  - By Ministries of Environment, Labor, Construction etc., LA by DANAK is referred in mandatory regulations including tests
  - "to be accredited by DANAK or equivalent EA
  - "equivalent" means EA-MLA members and ILAC-MRA members will be included in future
- Norway: NA
  - National Pollution Agency: measurement certificates for pollution and environment needs to be accredited by NA or MLA members
  - Tax reduction for environment: Tests to reduce tax, for example, sulfur measurement in exhaust, have to be accredited by NA or MLA member

However, many of regulators have not accepted accreditation in some economies, especially in big economies. This is the key issue for ILAC and each accreditation body.

10. Accreditation in legal metrology

In legal metrology, the authority needs to designate several organizations which have various roles in the domestic metrology system. An example for type approval is shown in Fig. 4, where dotted lines show authorization by the government. The purpose of the type approval is to confirm the quality of measuring instruments, shown by thick dotted line. For this purpose, technical criteria and/or guidelines are utilized by the legal metrology authority to designate organizations, NMIs, calibration and testing laboratories, certification bodies and manufacturers of measuring instruments according with the legal system.

At present, accreditation can cover all of these activities, using international standards and guidelines listed on Table 1. Therefore, it is possible to accredit or certify all players in type approval as shown by broken lines with arrows in Fig.4. If the testing laboratory tests the measuring instruments according to the OIML recommendations, as shown in italic characters in Fig. 4, all process could be covered by international criteria. If the legal metrology authority could accept the accreditation by an AB, the authority does not need to arrange their own technical activities for designation or authorization. Also, if the authority accepts oversea ABs or recognizes ILAC-MRA, it could be expected that the process to accept test results of foreign countries/economies became easy.

At present, the laboratory accreditation has been recognized among ABs with the ILAC MRA, mentioned in the previous section. However, this recognition is of among ABs not government. One of the most important targets of ABs and ILAC is to make accreditation and ILAC-MRA to be accepted by governments in various fields including legal metrology.

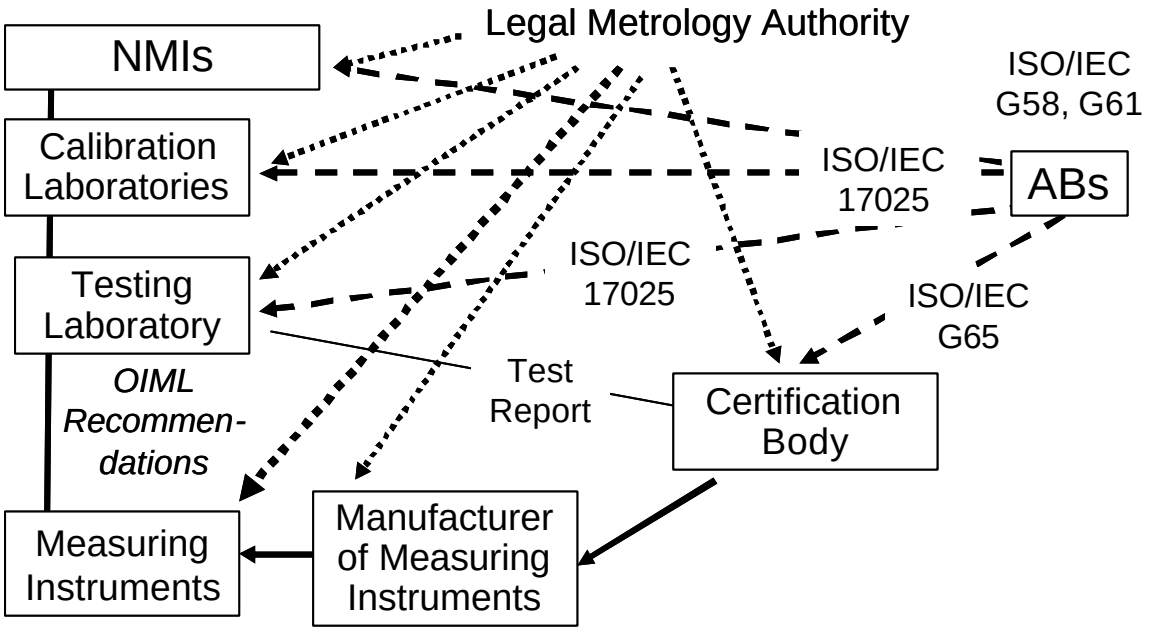


Fig.4 Traceability and Accreditation System for Type Approval

Uncertainty Analysis and Quality Control of Measuring Instruments

Akira Ooiwa

APLMF President
National Metrology Institute of Japan
National Institute of Advanced Industrial Science and Technology
AIST Tsukuba Central 3-1, 1-1-1 Umezono, Tsukuba, Ibaraki 305-8563, Japan

Abstract: There have been increasing demands for accurate measuring instruments with certain expression concerning its measurement confidence that should be available to other fields and also acceptable across borders. This strong trend has been caused by activities so called globalization which has been generating a lot of efforts of each economy to participate in the framework of the World Trade Organization (WTO).

There is an article in ISO-9000s as one of the quality management requirements, 'all the measurement results should have reasonable Traceability to achieve their confidence'. And also there is an definition for the word 'Traceability' such as ' property of the result of a measurement or the value of a standard whereby it can be related to stated references, usually national or international standards, through an unbroken chain of comparisons all having stated uncertainties'. It may be possible to imagine an idealistic situation of fully traceable measurements, but is quite difficult to realize an actual methodology that can achieve rational Traceability at reasonable cost and also can be acceptable to everyone as well. One of major difficulties to achieve such Traceability in fields is a lack of technical experience to apply the concept of 'uncertainty' to the field measurements operated with quality control method.

In this paper, for a typical measuring instruments of actual industrial application is examined, and a methodology are proposed to introduce 'uncertainty' into quality control of total measurement system in conjunction with already existing engineering techniques such as maximum permissible error (MPE) or tolerance, accuracy class of instruments, and regulation of instruments, etc.

Keywords:
Measuring instruments, traceability, uncertainty, accuracy class, tolerance, maximum permissible error (M.P.E.)

1.Requirements for Measurement in Fields

Nowadays it is ordinary that measurement results are displayed digitally and usually the result value is accurate enough with ɒ doubt as long as the measurement quality is properly controlled. In most of the cases it is true and there is technically no problem with an expression of accuracy or error, even if without expression of uncertainty.

There is a big difference between confidence and uncertainty in their conceptual meanings. The 'confidence' means a real total value of object, to be composed of quality and quantity. On the other hand the 'uncertainty' means a scientifically expressed numerical value of a certain physical condition. What is really wanted to know about the object by the person who is measuring it? The real purpose may be an evaluation result of the object with a certain confidence of quality that should satisfy the person. In order to treat the result objectively, it should be required to show a numerical value that has a scientific basis and with some expression of confidence of the value. An uncertainty itself does mean neither the quality nor the confidence of quality.

There have been happening confusions while introducing the idea of uncertainty into field measurements. One of the main reasons is mistaking 'uncertainty' for 'accuracy' or 'accuracy-class' that is similar to confidence in its meaning because it expresses a total performance of the measuring instrument. Although there is a difference in meanings between 'accuracy-class' (confidence) and 'uncertainty', many people have tried to simply change the words 'uncertainty' instead of 'accuracy'. Why it is so troublesome to introduce the concept of 'uncertainty' into engineering fields. The reason lies in the explanation of 'uncertainty' in the Guide of Expression of Uncertainty in Measurement (GUM). The meaning of 'uncertainty' is determined strictly scientifically to exclude the usage of 'accuracy' or 'error' those are the outcome of technical quality control and they are usually expressed and defined differently in each specific engineering field. The 'uncertainty' was introduced as a common concept to be used in all fields. A really needed action in order to use 'uncertainty' should be to try to describe the measurement confidence with using measurement result values accompanied by 'uncertainties'.

2.Quality Control of Measuring Instruments (MIs)

In engineering fields, all the measuring instruments should be maintained within certain acceptable condition, i.e. to achieve the traceability according to the regulation of ISO-9000s series. And the meaning of 'traceability' should be understood such that the measured value has to be related with an unbroken chain of comparisons to the reference standard representing SI units. Sometime people easily tend to think that comparison is equal to calibration. But the major part of engineering metrology consists of quality controls of both the measuring instrument and the condition of measurement, among which calibration is included as a part of them and is sometime a very small part. Therefore it is unreasonable to change or modify the total quality control system in order to introduce the idea of 'traceability with expression of uncertainty'. Instead the really necessary operation is to explain and prove the condition of existing quality control system by using the word 'uncertainty'. In this process it should be acceptable or recommended to explain the 'accuracy' and 'error' by using 'uncertainty' as long as they are efficiently applied in the actual system.

2.1 Quality Control Loops

In the engineering measurement there are usually quality control procedures with feedback loop as shown in Figure 1. One quality control loop consists basically of three steps and procedure flows. The first step is an acquisition of information, and the flow goes to the second step that is an evaluation of information based on some regulation. In case that the result of evaluation is good, the flow goes out of the loop with usable results. If the result of evaluation is bad, the flow goes to the third step that is to make adjustment or refinement. After the refinement, the process is looped back to the first step of information acquisition to restart the loop.

There are shown in the Figure 2 three quality control loops for Measuring Instruments quality system that are simplified but typical in the field application. Three loops are representing the quality control of MI at suppliers, the quality control of MI at MI user or maintenance, and the quality control of the measurement itself in fields, respectively. Through these quality controls it is achieved that the every measured value of the object should be reliable enough with a certain confidence level.

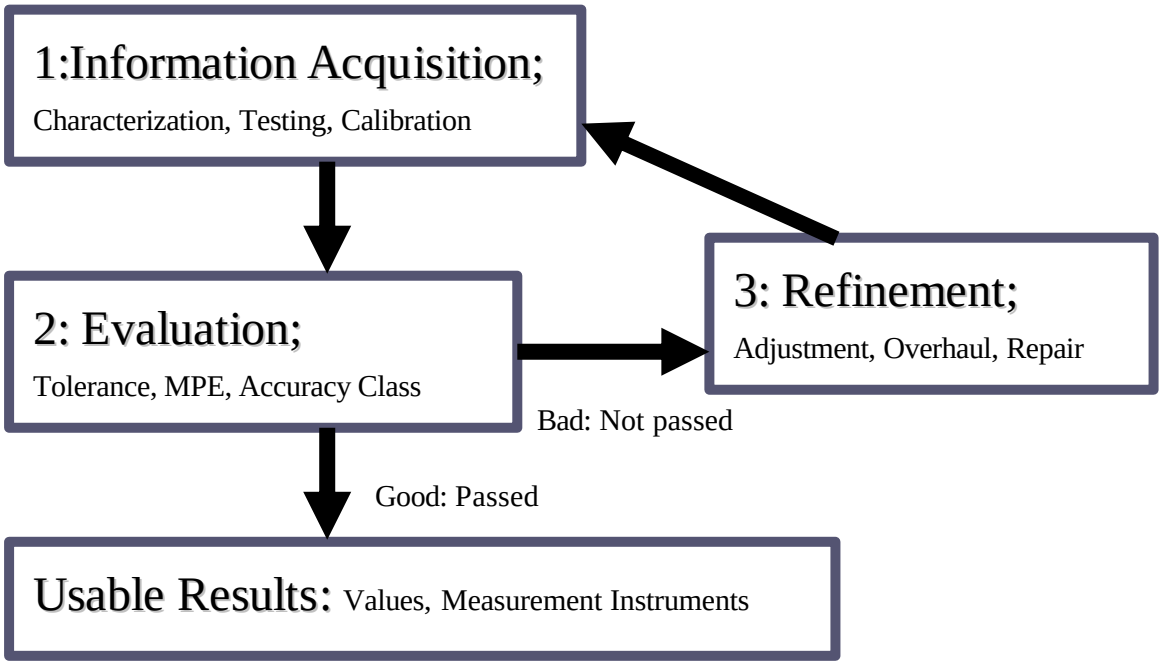


Figure1. Typical Loop of Quality Control

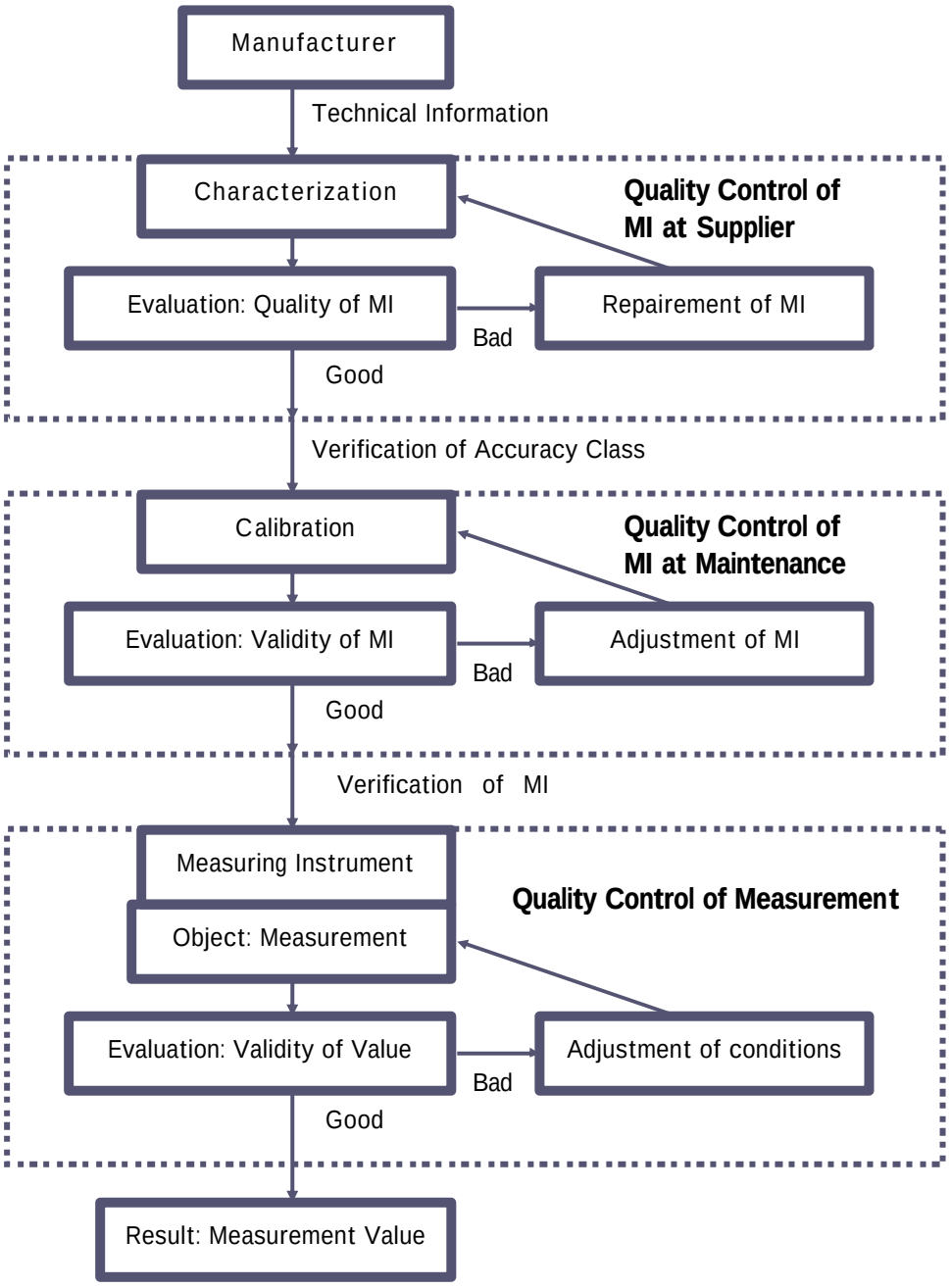


Figure 2. Schematic Flow Chart of Quality Control Loops of Measurements

2.2 Quality Control at Supplier

Using the information concerning MI those should be delivered by manufacturer, necessary fundamental characterizations and performance testing are carried out to verify that it has the expected performance. Main target performances should be concerning fundamental characteristics of measuring instruments such as indication linearity, hysteresis, resolution, sensitivities against environmental changes, durability against condition and time, and so on.

Calibration and adjustment may not be included in this process. Technical procedures for typical measuring instrument will be explained later. After characterization the evaluation should be done with comparing acquired characteristics results with regulation values those are usually tolerance values showing a acceptable range of the results. In quality control such tolerance method is efficient and usually applied. After the evaluation, when the results are not good, the MI should be repaired or adjusted, and transferred back to r re-characterization again as long as the refinement is successful. In case the result of evaluation is good against the target performance, the MI should be transferred to users in application fields. The outcome of this loop is a guarantee of performance of MI.

2.3 Quality Control at User/Maintenance

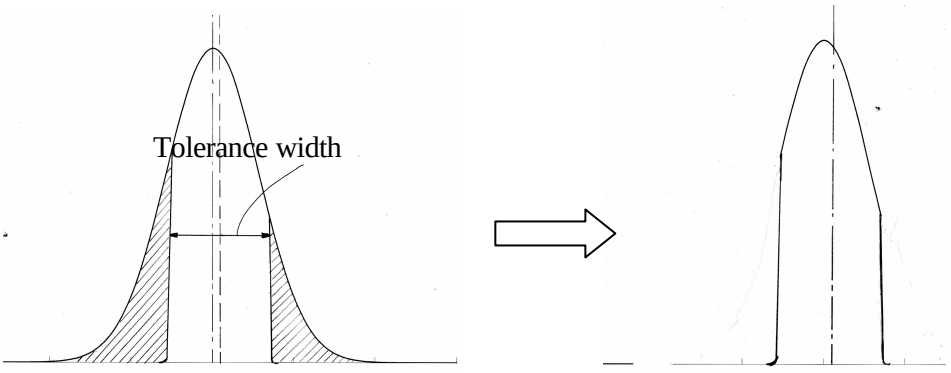
The MI should be calibrated periodically, and hopefully in situ condition. Calibration points are usually selected beforehand to find significant changes of the MI's performance efficiently. The main purpose of this process is to check whether the performance of the MI is maintained along with time. The important evaluation item is the amount of variation of calibrated value at each calibration point.

2.4 Quality Control of Measurement Results

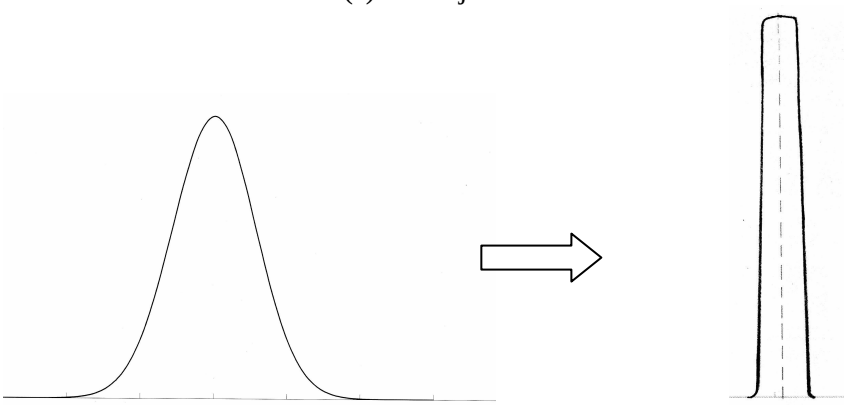
The MI is applied to an object to be measured. In this process evaluation should be done to make sure that the measured value represents the object itself using support data concerning circumstance and condition of the object. The final uncertainty of the measured value should be controlled within the tolerance that is calculated by statistical summation of all the effects of characteristics and duration time and condition change.

3. Quality Control Using the Idea of Tolerance or Maximum Permissible Error (MPE)

In almost all the cases of evaluation in the above mentioned quality control procedures, the idea of tolerance (or MPE) is applied. An ideal concept of tolerance method is shown in Figure 3. Supposing the original distribution of an object characteristic is approximately normal, and the uncertainty of the proving measurement is small enough comparing with the MI's distribution and the width of tolerance. When the value is out of the tolerance range, the MI should be eliminated and not be in use. Therefore the distribution of the characteristics of such passed MI becomes rectangular with a width same as of the tolerance. After this control, the possibility of getting out of the tolerance is very small and probable reasons are such as aging, breakdown of the MI, or an accident. Furthermore in case adjustment can be done, the distribution usually becomes of narrower rectangular shape than the width of tolerance according to the resolution of adjustment.



(a) No Adjustment



(b) With Adjustment

Figure 3. Reduction of Deviation by Quality Control Using Tolerance and Adjustment

4.Characterization of Measuring Instrument (MI)

Here is explained an example of characterization of a measuring instrument. Characteristics are classified to two categories, one is the fundamental characteristics such as repeatability, linearity, hysteresis, and resolution, those can be tested under some stable condition, and the other characteristics is concerning sensitivities of MI in relation with changes of circumstances such as temperature, atmospheric pressure and humidity, position and setting, shock and vibration, duration and aging, and so on. Typical methods to estimate these characteristics are explained in the followings sections.

4.1 Fundamental Characteristics: Characteristics under a Certain Standard Condition

Preparation: Before measurement, the Measuring Instrument should be set properly in stable conditions and principle procedures should be carried out according to the manufacturer's specifications such as,

a) Warm up the MI,
b) Set up the object to the MI, and
c) Make necessary check tests.

Characterization Test:

a) Choose several calibration points including the minimum and the maximum measurement points.

b) Determine the maintaining duration time to get stable indication of the MI from the data at the maximum and minimum measurement points.
c) Step up the input and measure the indication of the MI after keeping the maintaining duration time at each pressure point.
d) Keep the input at the maximum measurement point and saturate the input to be stable.

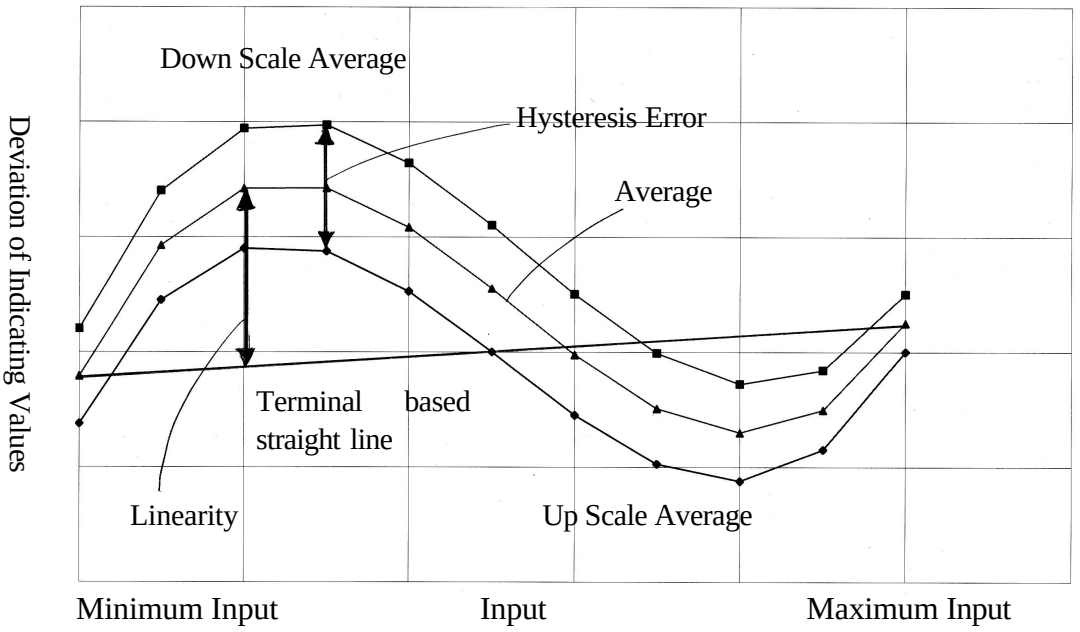


Figure 4. Typical Results of Fundamental Characteristics Testing

e) Step down the input and measure the indication of the MI same as in the process of stepping up.
f) Repeat from c) to e) procedures at least three times.

Estimation of the Characteristics of the MI Using the Data of Characterization Test:

Typical results of fundamental characteristics testing are shown in Fig.4. In this case indications are measured at every ten percent of the measurement range, and calibration curves are plotted for step up and down processes. Each point is averaged from three measurements. Using these data, specific characteristics are calculated according to the methods briefly explained as follows.

a) Linearity: the maximum deviation of averaged calibration curve from terminal based straight line.
b) Deviations at the minimum and maximum measuring values.
c) Hysteresis Error: the maximum difference between upscale and downscale average curves.
d) Repeatability: the maximum standard deviation of calibration points.
e) Resolution: the difference of input values that cause the smallest change of indication of the MI.

4.2 Sensitivity Characteristics Relating to Changes of Circumstances

Sensitivity characteristics estimations should be performed in case the indication of the measuring instrument varies with changes of conditions such as temperature and atmospheric pressure etc. Example methods to evaluate such characteristics are briefly described as follows.

a) Temperature characteristics: estimate from data of fundamental characteristics testing at various temperatures that include the minimum and the maximum temperatures that should be suggested in the specification of the MI by the manufacturer.
b) Power supply: estimate from fundamental characteristics testing at change input voltage within 10 % and also change input frequency within 5 %.
c) Position and setting: incline the MI and estimate the fundamental characteristics.
d) Vibration and shock: in case of necessity, estimate change before and after charging vibration or shock to the MI.
e) Long term stability: make testing for fundamental characteristics periodically or check simply the drift of zero point indication.

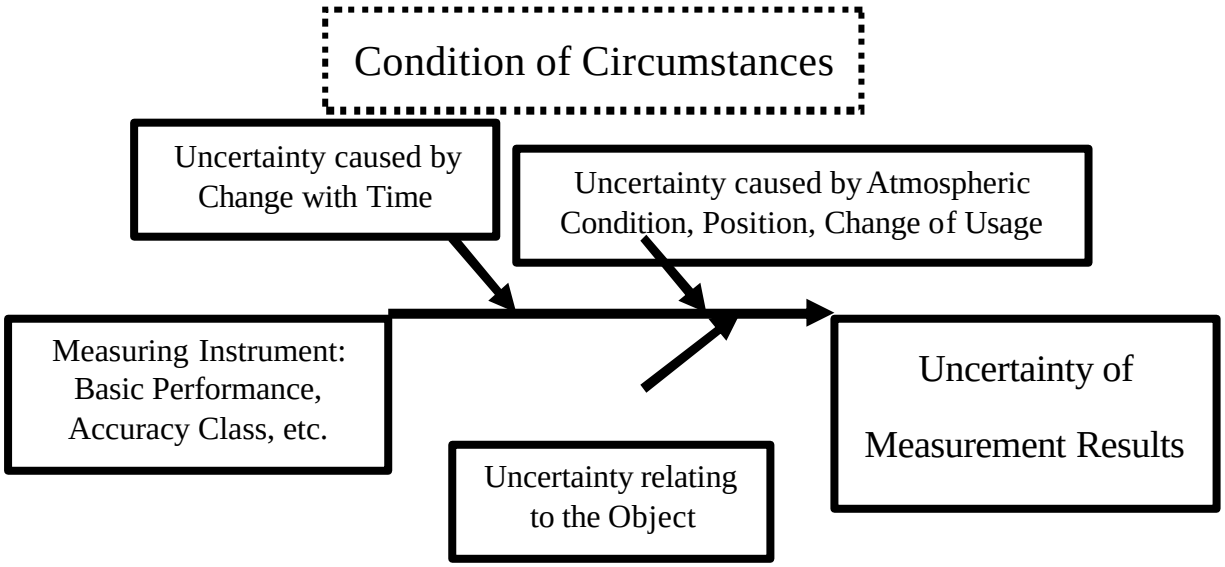


Fig. 5 Uncertainty of Measurement Results at Field

5. Uncertainty estimation of measurement value

There are three major factors for an uncertainty of the final measurement result value. The first factor is the performance of the MI to be estimated from the fundamental characteristics. The second factor is effects of conditions to the MI. The third factor is systematic relation between the MI and the object during the measurement.

5.1 Uncertainty concerning the performance of the MI

The uncertainty concerning the performance of the MI (*Up*) consists of Linearity, Deviations, Hysteresis, Repeatability, and Resolution. Data should be taken by the fundamental characteristics testing that is described in section 4.1. It is possible to easily calculate the uncertainty *Up* at each calibration point by making a statistical summation of these five elements. But a measuring value may not be same as that of calibration points, and it is not convenient to calculate *Up* at each new measuring value. Therefore, in order to simplify these calculations, a typical *Up* (*Upt*) is introduced which is derived by arithmetical summation of the maximum values of five factors over the whole measurement range. Because of this purpose calibration values should be selected properly from the measuring range.

5.2 Uncertainty concerning the change of condition

In case the circumstance conditions of the MI is not stable enough, the influence should be counted up for the uncertainty. In order to simplify calculations, every important condition is controlled within some specific range so called tolerance. Usually ambient temperature, humidity, pressure, and position are measured and controlled in certain tolerances those have been properly determined beforehand and the maximum deviation from the typical performance should be summed up with the *Upt*. As for such data for estimating the maximum possible change of performance may be provided by manufacturer. If one of the conditions changes bigger than the tolerance range, the measurement is not valid as long as such tolerance method is used.

As for the effect of duration or aging, it is difficult to estimate its real contribution to the uncertainty because it includes future expectations of the performance of the MI. Periodical checking or calibration would be a reasonable method to expect the possible maximum change of performance. Sometime manufacturer provide technical data about the duration of the MI. Possible maximum change with time should be counted and added to the uncertainty.

5.3 Uncertainty concerning object of measurement

The measured indication may not exactly represent the object state. In case the object condition is not stable then there should be some time delay in the indication. In case there are some disturbance or noise from other part of system, then there may be some shift of deviation in the indication. The user should investigate such possible factors and estimate the maximum influence to the indication, so as to add them to the uncertainty.

The final measurement value has an uncertainty composed of these three major components. Usually former two components are estimated using the idea of tolerance.

6. Conclusions

a) It is difficult to use 'uncertainty' directly in the quality control of MI, but it is recommended to use it indirectly with the aid of 'Tolerance' or 'Maximum Permissible Error (MPE)'.
b) Because there are many quality control loops concerning even one measurement, it is necessary to estimate all the loops and sum up all the factors of tolerances those affect the uncertainty of the final result.
c) It is essential to have technical information concerning characteristics of the MI in order to avoid duplicating testing and to perform cost effective measurement.
d) Manufacturer should provide such technical information and such information should be publicly transparent for all the possible users benefits.

7. Acknowledgment

The author wishes to thank the APEC secretariat for the support without which this report cannot be appeared, and the APLMF secretariat for all the organizations, and also thank participants who made valuable comments to this report at the symposium.

Uncertainty in Legal Metrology

Jong Oh Choi, Ph.D

Korea Research Institute of Standards and Science (KRISS)

P.O. Box 102, Yuseong Daejeon 305-600, Republic of Korea

Metrology plays a fundamental role in science, commerce and almost every aspects of human endeavor. Legal metrology provides a critical basis for the regulation of trade and the resolution of trade disputes. It has been recognized the measurement traceability is one of the key issues in various fields of measurements including legal metrology. As mentioned in the definition of traceability in VIM, as

"property of the results of a measurement or the value of a standard whereby it can be related to stated references, usually national or international standards, through an unbroken chain of comparisons all have stated uncertainties",

An establishment of traceability can be confirmed within the stated uncertainty. Therefore, the uncertainty and traceability become two essential factors for the quality of measurement result. Here, relating to the uncertainty in measurement and legal metrology, three topics are discussed.

1) Current problems of the term "accuracy"
   The definitions of "accuracy" in VIM and ISO are compared and discussed the problems occurred in the industrial sector due to the confliction of definition in legal metrology.
2) Calculation of Measurement Uncertainty based on GUM.
   The approach to calculate a measurement uncertainty based on the *"Guide to the Expression of Uncertainty of Measurement"* (GUM) is explained, which provides general rules for evaluating and expressing uncertainty in measurement across a broad spectrum of measurements [2].
3) Further Extension of GUM in repeated measurement
   The former approach is extended to determine the overall uncertainty by combining the uncertainties of the individual results of *n* measurements where the difference in the individual results is statistically significant. **Now, it is possible to determine the overall uncertainty by combining the uncertainties of the individual results *whether* the individual results are statistically different *or not*.**

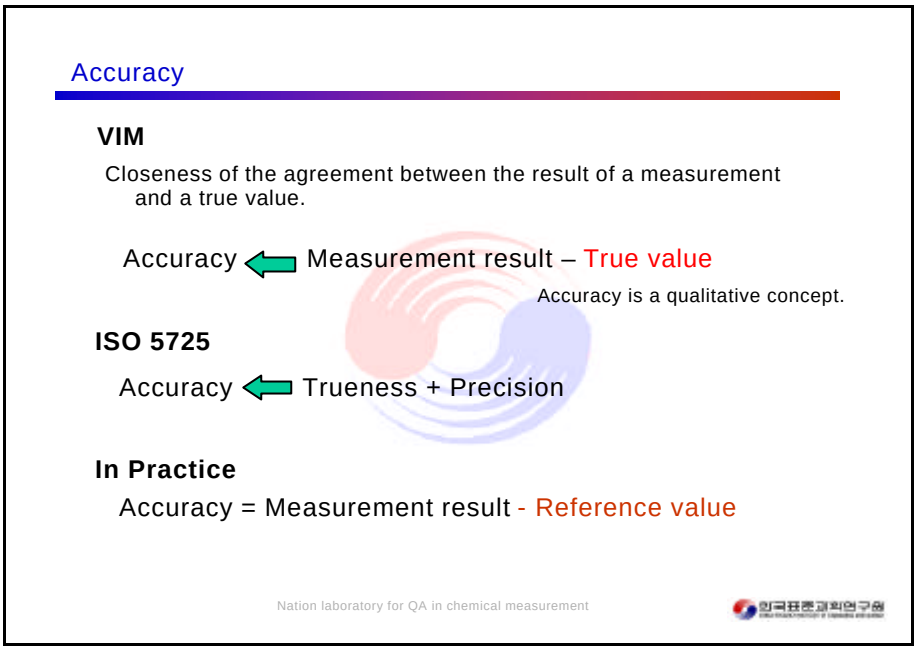
I. Current problems of the term “accuracy”

In order to discuss the measurement results, four basic terms, accuracy, measurand, quantity and uncertainty are described based on VIM.

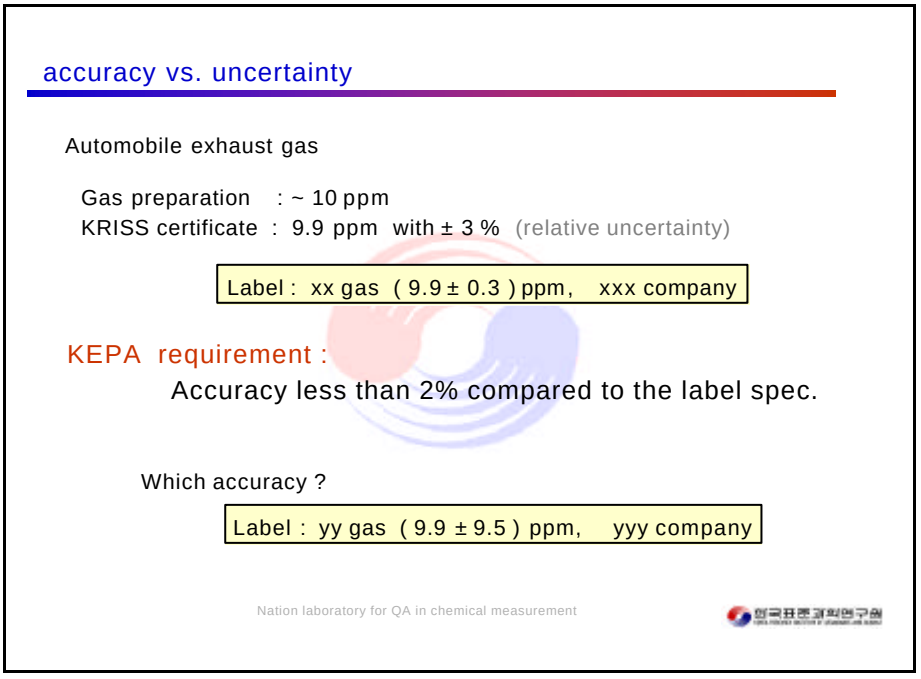
Accuracy is closeness of the agreement between the result of a measurement and a true value of the measurand. **Measurand** is a particular quantity subject to measurement. The specification of a measurand may require statements about quantities such as time, temperature and pressure. **Quantity** (measurable) is an attribute of a phenomenon, body or substance that may be distinguished qualitatively and determined quantitatively. **Uncertainty** is a parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurand.

On the contrary, ISO 5725 uses two terms “trueness” and “precision” to describe the accuracy of a measurement method. “Trueness” refers to the closeness of agreement between the arithmetic mean of a large number of test results and the true or accepted reference value. “Precision” refers to the closeness of agreement between test results. The need to consider precision arises because tests performed on presumably identical materials in presumably identical circumstances do not, in general, yield identical results. This is attributed to unavoidable random errors inherent in every measurement procedure; the factors that influence the outcome of a measurement cannot all be completely controlled. In the practical interpretation of measurement data, this variability has to be taken in to account. For instance, the difference between a test result and some specified value may be within the scope of unavoidable random errors, in which case a real deviation from such a specified value has not been established. Similarly, comparing test results from two batches of material will not indicate a fundamental quality difference if the difference between them can be attributed to the inherent variation in the measurement procedure. The general term accuracy is used in ISO 5725 to refer to both trueness and precision. Even though the term accuracy was at on time used to cover only the on component now named trueness, but it became clear that to many persons it should imply the total displacement of a result from a reference value, due to random as well as systematic effects.

In short, the following figure would be useful for the comparison.

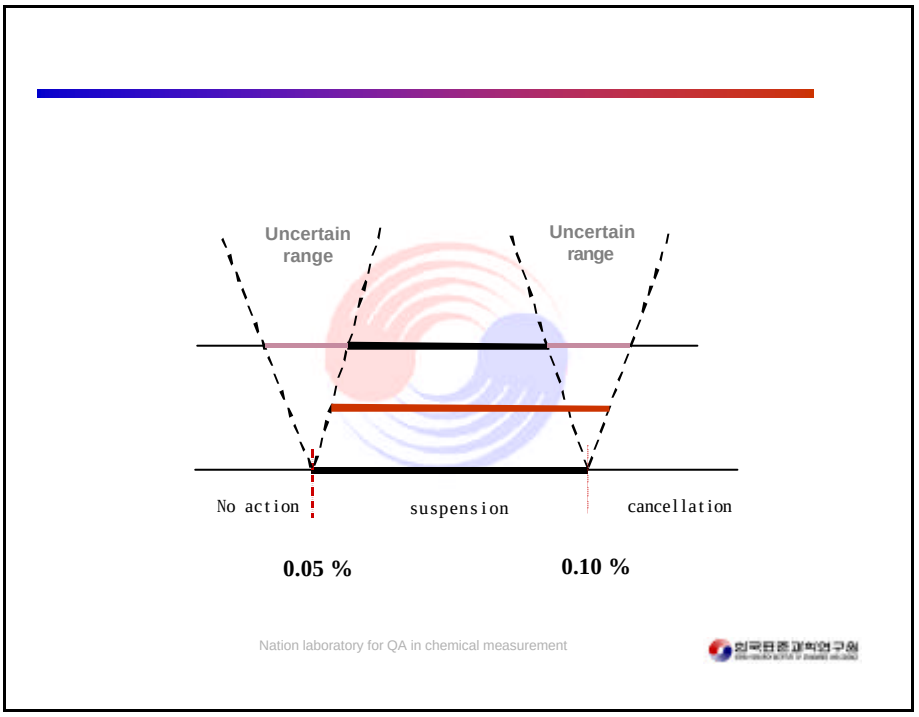


In VIM, accuracy is a qualitative concept, which can't be expressed in numbers, On the contrary, in practice; most legal documents employ reference value to calculate an accuracy. As is known, a measurement result indicates only the value itself at present. Accuracy and error are concerning about the value only. The following slide shows the problems of current traditional accuracy in legal metrology. As we know, accuracy and error are dealing with same magnitude in real application but different in the viewpoints.



In environmental legal metrology, it is requirement for a company to provide a reference gas for automobile exhaust gas measurement within 2% accuracy. A company prepared ~10 ppm gas, which is certified as 9.9 ppm with 3% relative expanded uncertainty by KRISS. Therefore, the company puts the result of 9.90 ± 0.03 ppm on the label.

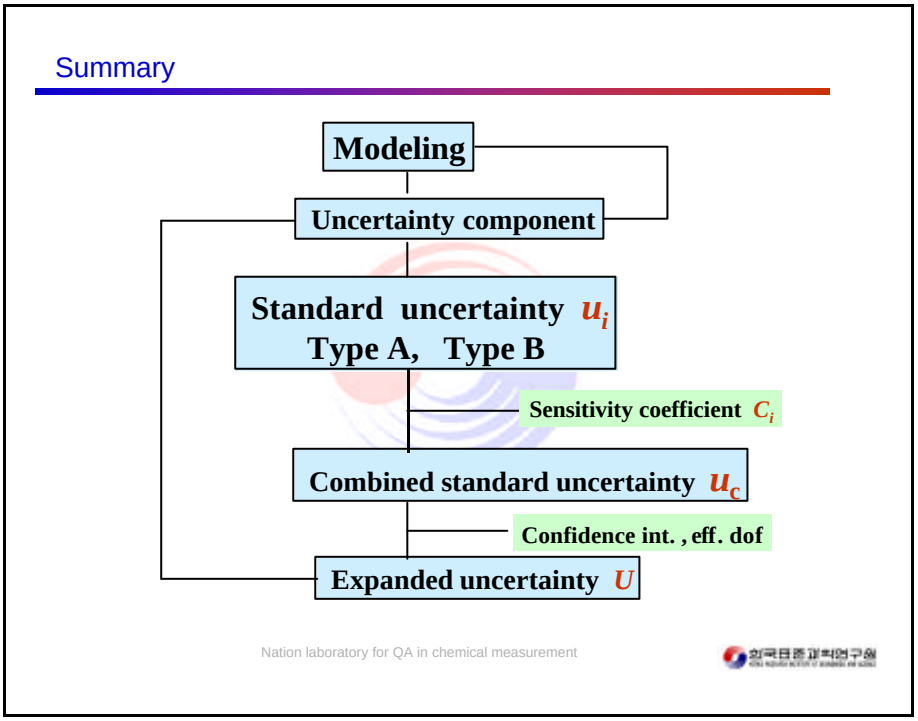
From a practical point of view for accuracy (difference between measurement result and reference value), the value 9.9 ppm has no error, since the measurement result and reference value are the same as 9.9 ppm. Considering the uncertainty of measurement results, it is quite obvious for the concentration of the product may not be within 2% accuracy (the difference between measurement result and true value). In other words, if we use the same criterion for practical accuracy for legal metrology, a gas product of 9.90 ± 9.5 ppm would be fine with the regulation. Now, it is the time for us to consider whether this is all right or not. The following is an example for a blood alcohol measurement in legal metrology and, with the consideration of the uncertainty; the range for conviction was modified in Korea in 2002.



II. Measurement Uncertainty by GUM

The *"Guide to the Expression of Uncertainty in Measurement"* (GUM) provides general rules for evaluating and expressing uncertainty in measurement across a broad spectrum of measurements [1]. The EURACHEM/CITAC Guide *"Quantifying Uncertainty in Analytical Measurement"* illustrates how the concepts in GUM can be applied in chemical measurement [2]. The examples in GUM and the EURACHEM guide are limited to a specific analytical determination employing *one* specific measurement procedure. Until now, most of the examples of uncertainty evaluation in diverse fields of measurements have been based on the result of *a single* measurement. However, it is quite common in analytical chemistry to carry out multiple measurements in order to report the result with an average value and its uncertainty. Recently, we have proposed an approach to uncertainty evaluation in which the overall uncertainty is determined by combining the uncertainties of the individual results in multiple measurements where the difference in the individual results is not statistically significant [3].

Here, we extended the approach to determine the overall uncertainty by combining the uncertainties of the individual results when the difference is statistically significant [4].



1. Basic definitions

1.1 Measurement equation

The case of interest is where the quantity Y being measured, called the **measurand**, is not measured directly, but is determined from N other quantities $X_1, X_2, \ldots, X_N$ through a functional relation f, often called the **measurement equation**:

$$Y = f(X_1, X_2, \ldots, X_N) \qquad (1)$$

Included among the quantities X_i are corrections (or correction factors), as well as quantities that take into account other sources of variability, such as different observers, instruments, samples, laboratories, and times at which observations are made (e.g., different days). Thus, the function f of equation (1) should express not simply a physical law but a measurement process, and in particular, it should contain all quantities that can contribute a significant uncertainty to the measurement result.

An estimate of the measurand or *output quantity* Y, denoted by y, is obtained from equation (1) using *input estimates* $x_1, x_2, \ldots, x_N$ for the values of the N *input quantities* $X_1, X_2, \ldots, X_N$. Thus, the *output estimate* y, which is the result of the measurement, is given by

$$y = f(x_1, x_2, \ldots, x_N) \qquad (2)$$

1.2 Meaning of uncertainty

If the probability distribution characterized by the measurement result y and its combined standard uncertainty $u_c(y)$ is approximately normal (Gaussian), and $u_c(y)$ is a reliable estimate of the standard deviation of y, then the interval $y \quad u_c(y)$ to $y + u_c(y)$ is expected to encompass approximately 68 % of the distribution of values that could reasonably be attributed to the value of the quantity Y of which y is an estimate. This implies that it is believed with an approximate level of confidence of 68 % that Y is greater than or equal to $y \quad u_c(y)$, and is less than or equal to $y + u_c(y)$, which is commonly written as $Y = y \pm u_c(y)$.

1.3 Classification of uncertainty components

The uncertainty of the measurement result y arises from the uncertainties $u(x_i)$ (or u_i for brevity) of the input estimates x_i that enter equation (2). Thus, in the example of equation (3), the uncertainty of the estimated value of the power P arises from the uncertainties of the

estimated values of the potential difference V, resistance R_0, temperature coefficient of resistance b, and temperature t. In general, components of uncertainty may be categorized according to the **method** used to evaluate them.

1. Type A evaluation
method of evaluation of uncertainty by the **statistical analysis** of series of observations,

2. Type B evaluation
method of evaluation of uncertainty by means **other than the statistical analysis** of series of observations.

1.4 Representation of uncertainty components

1. Standard Uncertainty
Each component of uncertainty, however evaluated, is represented by an estimated standard deviation, termed **standard uncertainty** with suggested symbol u_i, and equal to the positive square root of the estimated variance u_i^2.

2. Standard uncertainty: Type A
An uncertainty component obtained by a Type A evaluation is represented by a statistically estimated standard deviation s_i, equal to the positive square root of the statistically estimated variance s_i^2, and the associated number of degrees of freedom v_i. For such a component the standard uncertainty is $u_i = s_i$.

3. Standard uncertainty: Type B
In a similar manner, an uncertainty component obtained by a Type B evaluation is represented by a quantity u_j, which may be considered an approximation to the corresponding standard deviation; it is equal to the positive square root of u_j^2, which may be considered an approximation to the corresponding variance and which is obtained from an assumed probability distribution based on all the available information. Since the quantity u_j^2 is treated like a variance and u_j like a standard deviation, for such a component the standard uncertainty is simply u_j.

2. Evaluating uncertainty components:

2.1 Type A Evaluation

A Type A evaluation of standard uncertainty may be based on any valid statistical method for treating data. Examples are calculating the standard deviation of the mean of a series of independent observations; using the method of least squares to fit a curve to data in order to estimate the parameters of the curve and their standard deviations; and carrying out an analysis of variance (ANOVA) in order to identify and quantify random effects in certain kinds of measurements.

Mean and standard deviation

As an example of a Type A evaluation, consider an input quantity X_i whose value is estimated from n independent observations $X_{i,k}$ of X_i obtained under the same conditions of measurement. In this case the input estimate x_i is usually the **sample mean**

$$x_i = \bar{X}_i = \frac{1}{n}\sum_{k=1}^{n} X_{i,k} \qquad (4)$$

and the standard uncertainty $u(x_i)$ to be associated with x_i is the estimated **standard deviation of the mean**

$$u(x_i) = s(\bar{X}_i) = \left(\frac{1}{n(n-1)}\sum_{k=1}^{n}(X_{i,k} - \bar{X}_i)^2\right)^{1/2} \qquad (5)$$

2.2 Type B Evaluation

A Type B evaluation of standard uncertainty is usually based on scientific judgment using all of the relevant information available, which may include:

- previous measurement data,
- experience with, or general knowledge of, the behavior and property of relevant materials and instruments,
- manufacturer's specifications,
- data provided in calibration and other reports, and
- uncertainties assigned to reference data taken from handbooks.

Below are some examples of Type B evaluations in different situations, depending on the available information and the assumptions of the experimenter. Broadly speaking, the uncertainty is either obtained from an outside source, or obtained from an assumed distribution.

1. Uncertainty obtained from an outside source

Multiple of a standard deviation

Procedure: Convert an uncertainty quoted in a handbook, manufacturer's specification, calibration certificate, etc., that is a stated multiple of an estimated standard deviation to a standard uncertainty by dividing the quoted uncertainty by the multiplier.

Confidence interval

Procedure: Convert an uncertainty quoted in a handbook, manufacturer's specification, calibration certificate, etc., that defines a "confidence interval" having a stated level of confidence, such as 95 % or 99 %, to a standard uncertainty by treating the quoted uncertainty as if a normal probability distribution had been used to calculate it (unless otherwise indicated) and dividing it by the appropriate factor for such a distribution. These factors are 1.960 and 2.576 for the two levels of confidence given.

2. Uncertainty obtained from an assumed distribution

Normal distribution: "1 out of 2"

Procedure: Model the input quantity in question by a normal probability distribution and estimate lower and upper limits a_- and a_+ such that the best estimated value of the input quantity is $(a_+ + a_-)/2$ (i.e., the center of the limits) and there is 1 chance out of 2 (i.e., a 50 % probability) that the value of the quantity lies in the interval a_- to a_+. Then u_j is approximately 1.48 a, where $a = (a_+ - a_-)/2$ is the half-width of the interval.

Normal distribution: "2 out of 3"

Procedure: Model the input quantity in question by a normal probability distribution and estimate lower and upper limits a_- and a_+ such that the best estimated value of the input quantity is $(a_+ + a_-)/2$ (i.e., the center of the limits) and there are 2 chances out of 3 (i.e., a 67 % probability) that the value of the quantity lies in the interval a_- to a_+. Then u_j is approximately a, where $a = (a_+ - a_-)/2$ is the half-width of the interval.

Normal distribution: "99.73 %"

Procedure: If the quantity in question is modeled by a normal probability distribution, there are no finite limits that will contain 100 % of its possible values. However, plus and minus 3 standard deviations about the mean of a normal distribution corresponds to 99.73 % limits. Thus, if the limits a_- and a_+ of a normally distributed quantity with mean $(a_+ + a_-)/2$ are considered to contain "almost all" of the possible values of the quantity, that is, approximately 99.73 % of them, then u_j is approximately $a/3$, where $a = (a_+ - a_-)/2$ is the half-width of the interval.

Uniform (rectangular) distribution

Procedure: Estimate lower and upper limits a_- and a_+ for the value of the input quantity in question such that the probability that the value lies in the interval a_- and a_+ is, for all practical purposes, 100 %. Provided that there is no contradictory information, treat the quantity as if it is equally probable for its value to lie anywhere within the interval a_- to a_+; that is, model it by a uniform (i.e., rectangular) probability distribution. The best estimate of the value of the quantity is then $(a_+ + a_-)/2$ with $u_j = a$ divided by the square root of 3, where $a = (a_+ - a_-)/2$ is the half-width of the interval.

Triangular distribution

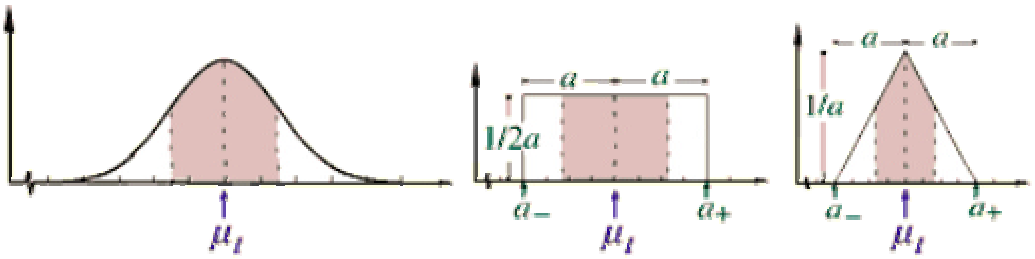
The rectangular distribution is a reasonable default model in the absence of any other information. But if it is known that values of the quantity in question near the center of the limits are more likely than values close to the limits, a normal distribution or, for simplicity, a triangular distribution, may be a better model.

Procedure: Estimate lower and upper limits a_- and a_+ for the value of the input quantity in question such that the probability that the value lies in the interval a_- to a_+ is, for all practical purposes, 100 %. Provided that there is no contradictory information, model the quantity by a triangular probability distribution. The best estimate of the value of the quantity is then $(a_+ + a_-)/2$ with $u_j = a$ divided by the square root of 6, where $a = (a_+ - a_-)/2$ is the half-width of the interval.

2.3 Schematic illustration of probability distributions

The following figure schematically illustrates the three distributions described above: normal, rectangular, and triangular. In the figures, μ_t is the expectation or mean of the distribution, and the shaded areas represent ± one standard uncertainty u about the mean. For a normal distribution, ± u encompasses about 68 % of the distribution; for a uniform

distribution, ± *u* encompasses about 58 % of the distribution; and for a triangular distribution, ± *u* encompasses about 65 % of the distribution.



3. Combining uncertainty components

3.1 Calculation of combined standard uncertainty

The **combined standard uncertainty** of the measurement result *y*, designated by $u_c(y)$ and taken to represent the estimated standard deviation of the result, is the positive square root of the estimated variance $u_c^2(y)$ obtained from

$$u_c^2(y) = \sum_{i=1}^{N} \left(\frac{\partial f}{\partial x_i} \right)^2 u^2(x_i) + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} u(x_i, x_j) \qquad (6)$$

Equation (6) is based on a first-order Taylor series approximation of the measurement equation $Y = f(X_1, X_2, \ldots, X_N)$ given in equation (1) and is conveniently referred to as the *law of propagation of uncertainty.* The partial derivatives of f with respect to the X_i (often referred to as *sensitivity coefficients*) are equal to the partial derivatives of f with respect to the X_i evaluated at $X_i = x_i$; $u(x_i)$ is the standard uncertainty associated with the input estimate x_i; and $u(x_i, x_j)$ is the estimated covariance associated with x_i and x_j.

Simplified forms

Equation (6) often reduces to a simple form in cases of practical interest. For example, if the input estimates x_i of the input quantities X_i can be assumed to be uncorrelated, then the second term vanishes. Further, if the input estimates are uncorrelated and the measurement equation is one of the following two forms, then equation (6) becomes simpler still.

Measurement equation:

A sum of quantities X_i multiplied by constants a_i.

$Y = a_1X_1 + a_2X_2 + \ldots a_NX_N$

Measurement result:

$y = a_1x_1 + a_2x_2 + \ldots a_Nx_N$

Combined standard uncertainty:

$u_c^2(y) = a_1^2u^2(x_1) + a_2^2u^2(x_2) + \ldots a_N^2u^2(x_N)$

Measurement equation:

A product of quantities X_i, raised to powers a, b, ... p, multiplied by a constant A.

$Y = AX_1^a X_2^b \ldots X_N^p$

Measurement result:

$y = Ax_1^a x_2^b \ldots x_N^p$

Combined standard uncertainty:

$u_{c,r}^2(y) = a^2u_r^2(x_1) + b^2u_r^2(x_2) + \ldots p^2u_r^2(x_N)$

Here $u_r(x_i)$ is the **relative standard uncertainty** of x_i and is defined by $u_r(x_i) = u(x_i)/|x_i|$, where $|x_i|$ is the absolute value of x_i and x_i is not equal to zero; and $u_{c,r}(y)$ is the **relative combined standard uncertainty** of y and is defined by $u_{c,r}(y) = u_c(y)/|y|$, where $|y|$ is the absolute value of y and y is not equal to zero.

4. Expanded uncertainty and coverage factor

4.1 Expanded uncertainty

Although the combined standard uncertainty u_c is used to express the uncertainty of many measurement results, for some commercial, industrial, and regulatory applications (e.g., when health and safety are concerned), what is often required is a measure of uncertainty that defines an interval about the measurement result y within which the value of the measurand Y can be confidently asserted to lie. The measure of uncertainty intended to meet this requirement is termed **expanded uncertainty**, suggested symbol U, and is obtained by multiplying $u_c(y)$ by a **coverage factor**, suggested symbol k. Thus $U = ku_c(y)$ and it is confidently believed that Y is greater than or equal to $y - U$, and is less than or equal to $y + U$, which is commonly written as $Y = y \pm U$.

4.2 Coverage factor

In general, the value of the coverage factor k is chosen on the basis of the desired level of confidence to be associated with the interval defined by $U = ku_c$. Typically, k is in the range 2 to 3. When the normal distribution applies and u_c is a reliable estimate of the standard deviation of y, $U = 2\ u_c$ (i.e., $k = 2$) defines an interval having a level of confidence of approximately 95 %, and $U = 3\ u_c$ (i.e., $k = 3$) defines an interval having a level of confidence greater than 99 %.

4.3 Relative expanded uncertainty

In analogy with relative standard uncertainty u_r and relative combined standard uncertainty $u_{c,r}$ defined above in connection with simplified forms of equation (6), the relative expanded uncertainty of a measurement result y is $U_r = U/|y|$, y not equal to zero.

III. Further Extension of GUM

Uncertainty evaluation based on GUM

A measurand Y is not measured directly, but is determined from N other input quantities X_i using a functional relationship f.

$$Y = f(X_1, X_2, \cdots, X_N) \tag{1}$$

where input quantities $X_1, X_2, \cdots, X_N$ upon which the output quantity Y depends may themselves be viewed as measurands and may themselves depend on other quantities.

The estimated standard deviation associated with each input estimate x_i of X_i is termed standard uncertainty $u(x_i)$. There are two approaches to estimate the standard uncertainty based on the evaluation method: Type A evaluation is based on statistical treatment on a series of observations, and Type B evaluation is based on all means other than statistical one such as data from certificates, manufacturer's specifications, previous experimental data, experience or knowledge, etc. For convenience, standard uncertainty estimated by Type A evaluation is sometimes called Type A standard uncertainty. The same applies for the Type B component. However, it is worthwhile considering that both Type A and Type B uncertainties can be due to a "random effect" as well as to a "systematic effect" in nature. Random effect gives rise to variations in repeated observations. Systematic effect is the recognized effect of an influence quantity on a measurement result, which causes systematic error. The estimated standard deviation associated with output estimate y of Y, termed combined standard uncertainty, $u_c(y)$, is:

$$u_c^2(y) = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i} \right)^2 u^2(x_i) + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left(\frac{\partial f}{\partial x_i} \right) \left(\frac{\partial f}{\partial x_j} \right) u(x_i, x_j) \tag{2}$$

where $u(x_i, x_j)$, covariance between x_i and x_j, is estimated by the degree of correlation

$$u(x_i, x_j) = r(x_i, x_j)\, u(x_i) u(x_j) \tag{3}$$

Although $u_c(y)$ can be universally used to express the uncertainty of a measurement result, it is often required to give a measure of uncertainty that defines an interval about the measurement result that may be expected to encompass a large fraction of the distribution of values that could reasonably be attributed to the measurand. This additional measure of uncertainty is termed the expanded uncertainty, U, and it is determined, in general, by multiplying $u_c(y)$ by a coverage factor $k=2$ [3].

$$U = k u_c(y) \tag{4}$$

An approach to combining uncertainties from multiple measurements

This approach is useful when the individual results are not statistically different. The individual results, y_1, y_2, …, y_n and their corresponding uncertainties, $u(y_1)$, $u(y_2)$, …, $u(y_n)$, are obtained by n measurements. The expected value m of M is taken as the arithmetic mean of n measurements.

$$M = \frac{1}{n}\sum_{k=1}^{n} Y_k = \frac{Y_1 + Y_2 + \cdots + Y_n}{n} \tag{5}$$

The combined standard uncertainty, based on Eq. (2), is

$$u_c^2(m) = \sum_{i=1}^{n}\left(\frac{\partial f}{\partial y_i}\right)^2 u^2(y_i) + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\left(\frac{\partial f}{\partial y_i}\right)\left(\frac{\partial f}{\partial y_j}\right)u(y_i, y_j) \tag{6}$$

According to GUM, an individual standard uncertainty $u(y_i)$ for the given single measurement has already been evaluated by combining the uncertainties from Types A and B evaluations. It can be recomposed into random $u_R(y_i)$ and systematic $u_S(y_i)$ components and be combined again to give

$$u(y_i) = \sqrt{u_R^2(y_i) + u_S^2(y_i)} \tag{7}$$

where $u_R(y_i)$ and $u_S(y_i)$ are the individual standard uncertainties due to random and systematic effects, respectively, of ith measurement. The separate combinations of random and systematic components for n measurements give the overall combined standard uncertainty

$$u_c(m) = \sqrt{u_R^2(m) + u_S^2(m)} \tag{8}$$

where $u_R(m)$ and $u_S(m)$ are the overall combined standard uncertainties due to random and systematic effects, respectively, for n measurements.

Since the individual uncertainties due to random effects in n measurements are not correlated, $r(y_i,y_j)=0$ in Eq. (6), the combined standard uncertainty due to random effects $u_R(m)$ is

$$u_R(m) = \sqrt{\frac{u_R^2(y_1) + u_R^2(y_2) + \cdots + u_R^2(y_n)}{n^2}} \tag{9}$$

On the contrary, the individual uncertainties due to systematic effects in n measurements are fully correlated, $r(y_i,y_j)=1$ in Eq. (6), therefore, the combined standard uncertainty due to systematic effects $u_S(m)$ is

$$u_S(m) = \frac{u_S(y_1) + u_S(y_2) + \cdots + u_S(y_n)}{n} \tag{10}$$

With the same measurement procedure under repeatability conditions, the uncertainties of the individual results are expected to be similar and, especially, the uncertainties due to systematic effects are supposed to be the same for n measurements. Therefore, Eq. (6) can be expressed as

$$u_C(m) = \sqrt{\frac{u_{R,P}^2(y)}{n} + u_S^2(y)} \tag{11}$$

where $$u_{R,P}(y) = \sqrt{\frac{u_R^2(y_1) + u_R^2(y_2) + \cdots + u_R^2(y_n)}{n}}$$

$$u_S(y) = u_S(y_1) = u_S(y_2) = \cdots = u_S(y_n)$$

An further extension of the previous extended approach

Given an estimate of difference in the individual results of n measurements, it is necessary to test whether it is statistically significant or not. If the significance test indicates no difference in the results, the previous approach will be applied. However, when the difference is found to be significant, it would be investigated in order to determine its cause and effect. If the difference is not due to a systematic effect, it results exclusively from a random effect.

In order to compensate the difference in the individual results, a new term $u_u^2(y)$ is introduced into Eq. (6) as an additional uncertainty resulting from a random effect, because the uncertainty arising from random effects can be reduced by n measurements [3].

$$u_c(m) = \sqrt{\frac{u_{R,P}^2(y) + u_u^2(y)}{n} + u_S^2(y)} \tag{7}$$

The term, $u_u^2(y)$, is supposed to be the between-group variance, $s_b^2(y)$, which has frequently been mentioned in the analysis of variance (ANOVA) methods [1].

$$u_u(y) = s_b(y) \tag{8}$$

In ANOVA, a variance in the individual results $s^2(y)$ is the measure of the sum of the between-group variance $s_b^2(y)$ and the within-group variance $s_w^2(y)/k$, where k is the number of measurements in the within-group.

$$s^2(y) = s_b^2(y) + \frac{s_w^2(y)}{k} \tag{9}$$

Therefore, the between-group variance in ANOVA is:

$$s_b^2(y) = s^2(y) - \frac{s_w^2(y)}{k} \tag{10}$$

In general, uncertainty evaluation requires information about systematic effects and random effects. The ANOVA techniques are designed to identify and evaluate the components of uncertainty arising from random effects. The term, $s_w(y)/\sqrt{k}$, is the standard uncertainty resulting from random effects, $u_{R,w}(y)$, which can be determined by k measurements in the within-group.

A well-performed measurement and proper evaluation of uncertainty arising from random effects will give

$$u_{R,P}(y) \cong u_{R,w}(y) = s_w(y)/\sqrt{k} \tag{11}$$

Thus, Eq. (10) can be written as

$$u_u^2(y) = s^2(y) - u_{R,P}^2(y) \tag{12}$$

Substituting Eq. (12) into Eq. (7) results in

$$u_c(m) = \sqrt{\frac{s^2(y)}{n} + u_S^2(y)} \tag{13}$$

where $s^2(y) = \sum_{i=1}^{n} (y_i - \bar{y})^2/(n-1)$

Discussion

The standard deviation of the individual results, $s(y)$, should reflect the standard uncertainty resulting from random effects, $u_{R,w}(y)$, in *a single* measurement. If the difference in the individual results of n measurements is statistically significant, it can be concluded that the standard deviation of the individual results is quite larger than the standard uncertainty resulting from random effects at the chosen level of significance. In other words, it is clear that the standard uncertainty resulting from random effects is underestimated in the evaluation.

In order to compensate the underestimation, an *unrecognized* uncertainty arising from a random effect, $u_u(y)$, is introduced into Eq. (6), which reflects the pure variation of the individual results without the measurement uncertainty arising from random effects. The effect, considered as an additional uncertainty, can be estimated by a type B evaluation in the corresponding measurement.

Now, by employing an *unrecognized* uncertainty arising from a random effect where the difference in the individual results is statistically significant, the difference becomes not significant any more. Therefore, Eq. (14) can be universally employed for uncertainty evaluation of *n* multiple measurements.

$$u_c(m) = \sqrt{\frac{u'^2_{R,P}(y)}{n} + u_S^2(y)} \qquad (14)$$

$$\text{where} \quad u'^2_{R,P}(y) = u^2_{R,P}(y) + u_u^2(y)$$

In addition, it is clear that Eq. (13) and Eq. (14) are the same since

$$s^2(y) = u^2_{R,P}(y) + u_u^2(y) \qquad (15)$$

In summary, if multiple measurements are well carried out and the uncertainty resulting from random effects is properly evaluated, $s(y)/\sqrt{k} = u_{R,w}(y) \cong u_{R,P}(y)$ is expected, since each term represents the uncertainty arising from random effects from different points of view. As a return, the consistency in those values can validate the uncertainty arising from random effects in the measurement.

Conclusions

Whether the individual results of *n* measurements are statistically different *or not*, it is possible to determine the overall uncertainty by combining the uncertainties of the individual results. The difference in the individual results is defined as an *unrecognized* uncertainty arising from a random effect. It is considered as an additional uncertainty arising from a random effect, which is estimated by a Type B evaluation. It is also shown that the uncertainty resulting from random effects can be confirmed by comparing the standard uncertainty of *a single* and/or *n* multiple measurement(s) to the standard deviation of the individual results of *n* measurements.

References

1. ISO GUM (1993) Guide to the Expression of Uncertainty in Measurement; ISO, Geneva, Switzerland
2. EURACHEM Guide (1995) Quantifying Uncertainty in Analytical Measurement; EURACHEM, Teddington, UK
3. JongOh Choi, Euijin Hwang, Hun-Young So, Byungjoo Kim (2003) Accred Qual Assur 8:13
4. JongOh Choi Dal-Ho Kim, Euijin Hwang, Hun-Young So (2003) Accred Qual Assur 8:205

Legal Metrology in Toy Testing for Safety

Foo-wing Lee, Yiu-chung Yip (Dr.), and Wai-hong Fung (Dr.)

Government Laboratory, Homantin Government Offices

88 Chung Hau Street, Homantin, Kowloon, Hong Kong China

Metrology

Metrology plays a fundamental role in science, commerce and almost every aspect of human endeavor. Most of the world's trade agreements now call for greater equivalency of measurement between trading partners, in which metrology plays an essential role. The major goal of such agreements is to eliminate double testing and other costly technical barriers to trade. In view of the recent proliferation of multilateral trade agreements, metrology has become a key facilitator of international trade.

Legal metrology provides a critical basis for the regulation of trade and the resolution of trade disputes. According to the *International vocabulary of terms in legal metrology* (VIML), legal metrology is defined as "the part of metrology relating to activities which result from statutory requirements and concern measurements, units of measurement, measuring instruments and methods of measurement and which are performed by competent bodies". A technical barrier to trade exists when a country applies technical regulations, standards or procedures for assessing conformity with these standards, in such a way as to impose an unnecessary restriction on international trade. The WTO Technical Barriers to Trade (TBT) Agreement goes some way towards addressing such barriers by requiring countries to act in a transparent and non-discriminatory manner. For this reason, the agreement encourages countries to use international standards where appropriate.

Toy Testing

Toy safety, being an aspect of human endeavor, is no exception to the above rule covered by legal metrology. An estimated hundreds of thousands of children are injured each year in accidents caused by unsafe toys. Although the primary objective of playing toys is to have fun, another very important aspect is that the playing should be safe and injury-free. It is thus necessary to ensure that toys are safe, which is always a prime concern of every government.

Hong Kong has a small domestic market for toys, be they locally produced or imported. Toys manufactured locally are largely for export. To avoid increasing business cost unnecessarily and to promote trade, the Hong Kong Government has decided not to introduce its own safety standards for toys. The local regulation, namely Toys and Children's Products Safety Ordinance, therefore adopts commonly accepted multiple national/international standards for compliance by toys available in Hong Kong. These standards have been introduced or used by our major trading partners as follows: International Voluntary Toy Safety Standard (IVTSS) established by the International Committee of Toy Industries, European Standard EN 71 established by the European Committee for Standardization, and ASTM F963 established by the American Society for Testing and Materials. The Ordinance provides that no person shall manufacture, import or supply a toy unless the toy complies with the applicable requirement contained in one of the above three sets of safety standards. In recent months, Hong Kong - China has been actively reviewing the Toys and Children's Products Safety Ordinance with a view to adopting the ISO 8124 series, together with ASTM F963 and EN-71, into the legal system. It is scheduled that the new legislation will come into effect in early 2004.

The ISO 8124 series consists of three parts, namely "ISO 8124-1:2000 Safety of Toys - Part 1: Safety Aspects Related to Mechanical and Physical Properties", "ISO 8124-2:1994 Flammability" and "ISO 8124-3:1997 Migration of Certain Elements". The requirements in ISO 8124 apply to all toys, i.e. any product or material designed or clearly intended for use in play by children under 14 years of age. They also apply after a toy is subjected to reasonably foreseeable conditions of normal use and abuse. The standard specifies acceptable criteria for structural characteristics of toys, such as shape, size, contour, spacing (e.g. rattles, small parts, sharp points and edges, hinge-line clearances) as well as acceptable criteria for properties peculiar to certain categories of toy (e.g. kinetic energy values for non-resilient-tipped projectiles, minimum tip angles for certain ride-on toys).

Laboratory Accreditation

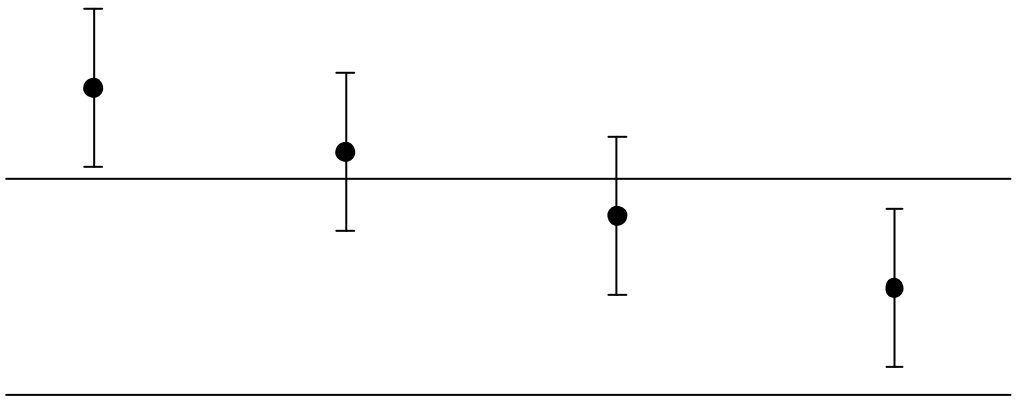
Laboratory accreditation is a formal recognition of technical competence of a testing laboratory for a specific task. In practice, it is a third party assessment based on the accreditation standard ISO/IEC 17025:1999[1]. The WTO has recognized non-acceptance of test results and measurement data as technical barrier to trade. Accreditation is considered to be the first essential step towards removing such technical barriers. For toy testing, merely following the ISO 8124 series is no guarantee for laboratories to provide fully compatible results in toy testing. Factors such as technical competency, equipment calibration, staff training, quality control checks, participating in proficiency testing may affect mutual recognition of test results generated among testing laboratories. Laboratory accreditation

based on ISO/IEC 17025 provides a harmonized means for showing technical competency of testing laboratories and establishing measurement traceability. For more than two years, the Hong Kong Government Laboratory has acquired accreditation for the ISO 8124 series through HOKLAS, the local laboratory accreditation body which has concluded mutual recognition agreements with over forty accreditation bodies in other economies.

Measurement Uncertainty

There is growing recognition of the importance of measurement uncertainty and traceability to assuring the reliability of measurements. Such importance is reflected in ISO/IEC 17025, which enhances requirements concerning the issue. In terms of compliance evaluation, it is not possible to make valid judgment without some knowledge of the traceability and measurement uncertainty. For toy testing, it is undesirable to make false positive decisions against safe products or false negative decisions for unsafe product. Measurement uncertainty thus plays an important role in making decisions, in particular when measured results are close to decision points of compliance. Two instances in compliance are clear-cut: either the results is above the upper limit, including its uncertainty, which means that the results is non-compliance (Figure 1a); or the results, including its uncertainty, is between the upper and lower limits, and is therefore, in compliance (Figure 1d). For any other case, some interpretation is necessary and can be made only in the light of the purpose, and with the knowledge and understanding of the end user of the information. Figure 1b represents probable noncompliance with the limit, while Figure 1c represents probable compliance. In both cases, noncompliance is not demonstrated beyond reasonable doubt.

Figure 1: Legal Compliance and Measurement Uncertainty



(a)	(b)	(c)	(d)
Result above limit plus uncertainty	Result above limit, but limit within uncertainty	Result below limit, but limit within uncertainty	Result below limit minus uncertainty

A number of ways are available for estimation the uncertainty of a measurement system. The approach needed to deal with traceability and measurement uncertainty depends on the intended use of the testing results. The best practice in evaluating measurement uncertainty is described in ISO Guide to the Expression of Uncertainty in Measurement[2] (hereafter referred to as GUM). Varying degrees of rigor and sophistication can be used in the evaluation of measurement uncertainty ranging from full and costly evaluation of all sources of uncertainty to rough and ready estimated based on available information.

According to GUM, it is a general rule to understand the measurement process before making a measurement, otherwise inappropriate measurements may be made. Once understood it is appropriate to make a model of the system which may be represented in the form of an equation, formula or diagram before commencing calculations. The uncertainty of a test result is a combination of a number of uncertainty components. Even a single instrument reading may be influenced by several factors. Careful consideration of each step involved in the test is required to identify and list all the factors that contribute to the uncertainty. This is the most crucial stage and requires a good understanding of the measurement equipment, the principles and practice of the test and the influence of environment. The next step is to quantify uncertainty components by appropriate means. An initial approximate quantification may be valuable in enabling some components to be shown to be negligible and not worthy of more rigorous evaluation. In most cases a practical definition of negligible would be a component that is not more than a fifth of the size of the largest component. Some components may be quantified by calculation of the standard deviation from a set of repeated measurements. Quantification of others will require the exercise of judgment using all relevant information on the possible variability of each factor, including: (a) previous measurement data; (b) manufacturer's specifications; (c) data provided in calibration certificates; (d) uncertainty assigned to reference data taken from handbooks; (e) experience with or general knowledge of the behaviour and properties of relevant materials and instruments; evaluations made under this heading are quite common in many fields of testing, but must be made with care and by suitably experienced personnel; (f) results of an interlaboratory comparison test programme. The standard uncertainty is expressed as the standard deviation. The standard uncertainty components are combined to produce the combined standard uncertainty of the result. It is usually necessary to quote an 'expanded uncertainty', and the combined standard uncertainty therefore needs to be multiplied by the appropriate 'coverage factor'. This must reflect the level of confidence required. In general, a value of 2 for the coverage factor can be taken to define an interval having a level of confidence of approximately 95%.

Estimation of Uncertainty for Selected tests of ISO 8124

There are more than 100 separate tests and design specifications prescribed in the ISO 8124 series to reduce or eliminate hazards with the potential to cause injury under conditions of normal use or reasonably foreseeable abuse. These include testing for accessible sharp points and edges, measuring for small parts, wheel-pull resistance and projectiles, testing for flammability, toxicity and thermal requirements, as well as acoustical requirements. The following examples illustrate our practical experience in estimating measurement uncertainty of selected tests stipulated in the ISO 8124 series.

Example 1: toys containing a heat source (Clause 4.23, ISO 8124-1:2000)

The objective of the clause is to prevent burning hazard from toys containing a heat source. The testing procedures, as given in Clause 5.18 of the standard, are relatively straightforward. In brief, the toy is operated in accordance with the instructions for use at the maximum input until equilibrium temperature is reached in an ambient temperature of 20°C±5°C. Temperatures of accessible parts of the toy are measured by thermocouples before and after the operation and any temperature rises are calculated by the temperature differences. The model for uncertainty analysis is given below:

Temperature rise, y (°C) = final temperature, y_{final} (°C) – initial temperature, $y_{initial}$ (°C)

Based on this model, the uncertainty budget in calculating $U(y)$ shall include the budgets in the measurement of $y_{initial}$ and y_{final}.

(1) The initial temperature ($y_{initial}$) measured at the selected accessible part of the toy is obtained from the equation:

$$y_{initial} = \bar{y}_{initial} + A + \delta B$$

Three are three components of uncertainty in the initial temperature reading, one arising from the variation in repeated readings ($\bar{y}_{initial}$), the other from the variations in correction to thermometer reading (A), and the last from variations in rounding of the value of the least significant digit of the thermocouple (δB). The uncertainty budget for initial temperature can be summarised as follows:

Symbol	Uncertainty/ Semi-range, (°C)	Type	Probability distribution	Divisor	Standard uncertainty, $u(y)$	Degrees of freedom, v_i
A	a	B	Normal	k_a	$\frac{a}{k_a}$	v_a
δB	b	B	Rectangular	$\sqrt{3}$	$\frac{b}{\sqrt{3}}$	Infinite
$\bar{y}_{initial}$	c	A	Normal	$\sqrt{n}$	$\frac{c}{\sqrt{n}}$	$n-1$
Combined standard uncertainty, $u(y_{initial})$					$\sqrt{\sum_{i=1}^{N} u_i^2(y)}$	
Effective degrees of freedom, $v(y_{initial})$						$\frac{u^4(y_{initial})}{\sum_{i=1}^{N} \frac{u_i^4(y)}{v_i}}$

(2) The final temperature (y_{final}) measured at the selected accessible part of the toy is obtained from:

$$y_{final} = \bar{y}_{final} + D + \delta B$$

Similar to initial temperature measurement, three components of uncertainty are included in the final temperature reading, one arising from the variation in repeated readings ($\bar{y}_{final}$), the other from the variations in correction to thermometer reading (D), and the last from variations in rounding of the value of the least significant digit of the thermocouple (δB). The uncertainty budget for final temperature can be summarised as follows:

Symbol	Uncertainty/ Semi-range, (°C)	Type	Probability distribution	Divisor	Standard uncertainty, $u(y)$	Degrees of freedom, v_i
D	d	B	Normal	k_d	$\frac{d}{k_d}$	v_d
δB	b	B	Rectangular	$\sqrt{3}$	$\frac{b}{\sqrt{3}}$	Infinite
$\bar{y}_{final}$	e	A	Normal	$\sqrt{n}$	$\frac{e}{\sqrt{n}}$	$n-1$
Combined standard uncertainty, $u(y_{final})$					$\sqrt{\sum_{i=1}^{N} u_i^2(y)}$	
Effective degrees of freedom, $v(y_{final})$						$\frac{u^4(y_{final})}{\sum_{i=1}^{N} \frac{u_i^4(y)}{v_i}}$

(3) The expanded combined uncertainty $U(y)$, based on ISO GUM, is as follows:

Value (°C)	Standard uncertainty, $u_i(y)$	Degrees of freedom, v_i
$y_{initial}$	$u(y_{initial})$	$v(y_{initial})$
y_{final}	$u(y_{final})$	$v(y_{final})$
Combined standard uncertainty, $u_c(y)$	$\sqrt{\sum_{i=1}^{N} u_i^2(y)}$	
Effective degrees of freedom, $v(y)$		$\frac{u_c^4(y)}{\sum_{i=1}^{N} \frac{u_i^4(y)}{v_i}}$
Expanded uncertainty, $U(y)$	$k\ u_c(y)$ (Note: the effective degrees of freedom v_{eff} and table G.2 in the GUM are used to determine the value of coverage factor k for a level of confidence of 95 %.)	

Example 2: kinetic energy of projectile toys (Clause 4.18, ISO 8124-1:2000)

Projectile toys store with kinetic energy. To prevent potential hazards that might be caused by projectile toys and by the firing of improvised projectiles from such toys, it is necessary that the maximum kinetic energy released by these toys under normal use should be under control. Testing procedures for measuring kinetic energy are described in Clause 5.15 of ISO 8124-1:2000. In practice, the mass of the projectile toy is measured with an analytical balance, the time taken by the projectile to travel a known distance is measured by a chronograph, while the distance traveled by the projectile is determined by a steel rule. The projectile velocity can then be obtained according to the equation v = d/t, which is used for calculation of the kinetic energy of the toy. The model for uncertainty analysis is given below:

$$\text{K.E.} = \frac{1}{2}mv^2$$

$$= \frac{1}{2}m\left(\frac{d}{t}\right)^2$$

$$= \frac{1}{2}m\frac{d^2}{t^2}$$

where, K.E. = kinetic energy (J);

m = mass of the projectile (kg);

v = velocity of projectile (ms^{-1}) = d/t;

d = distance between the sensors of the chronograph (m);

t = time (s) taken by the projectile to travel the known distance d

Based on the above model, the uncertainty budget in calculating U(K.E.) includes the uncertainty associated with measurement of projectile mass, projectile distance, and projectile traveling time.

(1) The mass of the projectile is obtained from:

$$\mathrm{m} \qquad \overline{\mathrm{m}} + \delta A + \delta B + \delta C$$

There are four components of uncertainty for measurement of projectile mass (m), one arising from the variation in repeated readings ($\overline{\mathrm{m}}$), the second from variations in linearity of analytical balance (δA), and the third from variations in rounding of the value of the least significant digit of the balance (δB), and the fourth from variations in correction for air buoyancy (δC). To simplify the model, the variation due to air buoyancy correction is not included in the budget because all weighing results are quoted on the conventional basis for weighing in air. The overall uncertainty budget for projectile mass is summarised as follows:

Symbol	Uncertainty/ Semi-range, (kg)	Type	Probability distribution	Divisor	Standard uncertainty, $u(x_i)$	Degrees of freedom, v_i
δA	a	B	Normal	$\sqrt{3}$	$\sqrt{2\times\left(\frac{a}{\sqrt{3}}\right)^2}$	Infinite
δB	b	B	Rectangular	$\sqrt{3}$	$\frac{b}{\sqrt{3}}$	Infinite
$\overline{\mathrm{m}}$	d	A	Normal	$\sqrt{n}$	$\frac{d}{\sqrt{n}}$	$n-1$
Combined standard uncertainty, u(m)					$\sqrt{\sum_{i=1}^{N} u(x_i)^2}$	
Effective degrees of freedom, v(m)						$\frac{u(\mathrm{m})^4}{\sum_{i=1}^{N}\frac{u(x_i)^4}{v_i}}$

(2) The projectile distance between the sensors of the chronograph is obtained from:

$$\mathrm{d} \quad = \quad \mathrm{d}' + \delta E + \delta F$$

Three components of uncertainty are included in the measurement of projectile distance (d), one arising from the variation in repeated readings (d′), the second from variations in calibration of the steel rule (δE), and the third from variations in rounding of the value of nearest division of the steel rule (δF). The uncertainty budget is summarised below:

Symbol	Uncertainty/ Semi-range, (m)	Type	Probability distribution	Divisor	Standard uncertainty, $u(x_i)$	Degrees of freedom, v_i
δE	e	B	Normal	k_e	$\frac{e}{k_e}$	v_e
δF	f	B	Rectangular	$\sqrt{3}$	$\frac{f}{\sqrt{3}}$	Infinite
d′	g	A	Normal	$\sqrt{n}$	$\frac{g}{\sqrt{n}}$	$n-1$
Combined standard uncertainty, u(d)					$\sqrt{\sum_{i=1}^{N} u(x_i)^2}$	
Effective degrees of freedom, v(d)						$\frac{u(\mathrm{d})^4}{\sum_{i=1}^{N} \frac{u(x_i)^4}{v_i}}$

(3) The time taken by the projectile to travel the known distance d is obtained from:

$$\mathrm{t} = \mathrm{t}' + H + \delta J$$

As shown in the above, three components of uncertainty are included in the measurement of projectile time (t), one arising from the variation in repeated time measurements taken by the projectile to travel the known distance d as recorded by the chronograph (t'), the other from the variations in correction of the chronograph (H), and the last from variations in rounding of the value of the least significant digit of the chronograph (δJ). Since the testing method requires only the shortest projectile time rather than averaged time of repeated analysis for reporting, t' is thus a constant and is not included in the budget. The uncertainty budget for the time taken by the projectile to travel the known distance, d, is summarised below:

Symbol	Uncertainty/ Semi-range, (s)	Type	Probability distribution	Divisor	Standard uncertainty, $u(x_i)$	Degrees of freedom, v_i
H	h	B	Normal	k_h	$\frac{h}{k_h}$	v_h
δJ	j	B	Rectangular	$\sqrt{3}$	$\frac{j}{\sqrt{3}}$	Infinite
Combined standard uncertainty, u(t)					$\sqrt{\sum_{i=1}^{N} u(x_i)^2}$	
Effective degrees of freedom, v(t)						$\frac{u(\mathrm{t})^4}{\sum_{i=1}^{N} \frac{u(x_i)^4}{v_i}}$

(4) Since the model adopted in this example is not based on simple arithmetic, it is not appropriate to assume the sensitivity coefficient as unity. We have initially calculated the expanded combined uncertainty of U(K.E.) based on partial differential approach as described in GUM. However we noticed that an equivalent result on expanded uncertainty can also be obtained when the calculation is done by spreadsheet approach. To avoid complicated mathematics in daily work, the spreadsheet approach is more desirable. Details of the two approaches are summarised below:

(i) Calculation of U(K.E.) by spreadsheet approach:

Value, x_i	Standard uncertainty, $u(x_i)$	Relative standard uncertainty, $u(x_i)/x_i$	Degrees of freedom, v_i
M (kg)	u(m)	$\frac{u(\mathrm{m})}{\mathrm{m}}$	v(m)
D (m)	u(d)	$\frac{u(\mathrm{d})}{\mathrm{d}}$	v(d)
t (s)	u(t)	$\frac{u(\mathrm{t})}{\mathrm{t}}$	v(t)
Combined standard uncertainty, u_c(K.E.)	$\mathrm{K.E.} \times \sqrt{\left(\frac{u(\mathrm{m})}{\mathrm{m}}\right)^2 + \left(\frac{2\,u(\mathrm{d})}{\mathrm{d}}\right)^2 + \left(\frac{2\,u(\mathrm{t})}{\mathrm{t}}\right)^2}$		
Effective degrees of freedom, v(K.E.)			$\frac{[u_c(\mathrm{K.E.})/\mathrm{K.E.}]^4}{\frac{\left(\frac{u(\mathrm{m})}{\mathrm{m}}\right)^4}{v(\mathrm{m})} + \frac{\left(\frac{2\,u(\mathrm{d})}{\mathrm{d}}\right)^4}{v(\mathrm{d})} + \frac{\left(\frac{2\,u(\mathrm{t})}{\mathrm{t}}\right)^4}{v(\mathrm{t})}}$
Expanded uncertainty, U(K.E.)	$k\,u_c$(K.E.) (Note: the effective degrees of freedom v_{eff} and table G.2 in the GUM are used to determine the value of coverage factor k for a level of confidence of 95%.)		

(ii) Calculation of U(K.E.) by partial differentia l approach:

Value, x_i	Standard uncertainty, $u(x_i)$	Sensitivity coefficient, c_i	Degrees of freedom, v_i
m (kg)	u(m)	$\frac{\partial(\text{K.E.})}{\partial \text{m}} = \frac{\text{d}^2}{2\text{t}^2}$	v(m)
d (m)	u(d)	$\frac{\partial(\text{K.E.})}{\partial \text{d}} = \frac{\text{md}}{\text{t}^2}$	v(d)
t (s)	u(t)	$\frac{\partial(\text{K.E.})}{\partial \text{t}} = -\frac{\text{md}^2}{\text{t}^3}$	v(t)
Combined standard uncertainty, u_c(K.E.)	$\sqrt{\sum_{i=1}^{N}[c_i\, u(x_i)]^2}$		
Effective degrees of freedom, v(K.E.)			$\frac{u_c^4(\text{K.E.})}{\sum_{i=1}^{N}\frac{[c_i\, u(x_i)]^4}{v_i}}$
Expanded uncertainty, U(K.E.)	$k\, u_c$(K.E.) (Note: the effective degrees of freedom v_{eff} and table G. 2 in the GUM are used to determine the value of coverage factor k for a level of confidence of 95 %.)		

Example 3: Sound pressure level (Annex F.3, ISO 8124-1:2000)

To protect young children against loud noise generated by sound-emitting toys, the ISO 8120 provides in its Annex F.3 an informative requirement for the maximum sound pressure levels produced by such toys. Determination of the A-weighted emission sound pressure level (LpA) for *close-to-the-ear toy* is given as an example here. The measurement makes use of a sound pressure level meter at a specified distance of 2.5 cm ± 0.5 cm from that surface of the toy where the main sound source exists. The uncertainty model consists of six components:

$$y_1 = y_1' + \delta A + \delta B + \delta C + \delta D + \delta E$$

where,

y_1' = The highest value recorded (L_{pA}) at any of the microphone positions;

δA = The rounding of the value of the least significant digit of the sound level meter;

δB = Calibration of the sound level calibrator;

δC = Drift of the sound level calibrator since last calibration;

δD = Tolerances for the pulse range; and

δE = Tolerances on frequency weighting characteristics.

The uncertainty budget for A-weighted emission sound pressure level at the specified position, L_{pA}, is summarized below:

Symbol	Uncertainty/ Semi-range, (dB)	Type	Probability distribution	Divisor	Standard uncertainty, $u(y)$	Degrees of freedom, v_i
δA	a	B	Normal	$\sqrt{3}$	$\frac{a}{\sqrt{3}}$	Infinite
δB	b	B	Normal	k_b	$\frac{b}{k_b}$	v_b
δC	c	B	Rectangular	$\sqrt{3}$	$\frac{c}{\sqrt{3}}$	Infinite
δD	d	B	Rectangular	$\sqrt{3}$	$\frac{d}{\sqrt{3}}$	Infinite
δE	e	B	Rectangular	$\sqrt{3}$	$\frac{e}{\sqrt{3}}$	Infinite
y_1'	f	A	Normal	$\sqrt{n}$	$\frac{f}{\sqrt{n}}$	$n - 1$
Combined standard uncertainty, $u_c(y)$					$\sqrt{\sum_{i=1}^{N} u_i^2(y)}$	
Effective degrees of freedom, $v(y_1)$						$\frac{u_c^4(y_1)}{\sum_{i=1}^{N} \frac{u_i^4(y)}{v_i}}$
Expanded uncertainty, $U(y)$	$k\ u_c(y)$ (Note: the effective degrees of freedom v_{eff} and table G. 2 in the GUM are used to determine the value of coverage factor k for a level of confidence of 95%.)					

where,

a: half least significant digit to which the sound level meter responds on the calibrated range.

b: the expanded uncertainty as stated in the certificate of calibration of the sound level calibrator for a level of confidence of 95% with degrees of freedom v_b and a coverage factor k_b.

c: no correction is made for the drift of the nominal pressure field calibration signal of the sound level calibrator. The limits of $\pm\ c$ are estimated from the results of the calibration history of the sound level calibrator. The probability distribution is assumed to be rectangular.

d: Table 2 of IEC 60804:1985[3] showing the minimum values for pulse range with tolerances is extracted as follows:

Tolerance, in decibels	Type 0	Type 1
Minimum value for pulse range	73	63
Tolerance, burst duration < 10 ms but ≥ 1 ms	± 1.9	± 2.2
Tolerance, burst duration ≥ 10 ms	± 1.4	± 1.7

The tolerances for the pulse range are estimated to have limits of ± *d*. The probability distribution is assumed to be rectangular.

e: the certificate of calibration of the sound level meter indicated that the deviations between the actual SLM readings and the expected SLM readings for the A- and C-frequency weightings were found to be within ± 0.2dB. The frequency weighting characteristics are estimated to have limits of ± *e*. The probability distribution is assumed to be rectangular. Table 5 of IEC 60651:1979[4] showing tolerances on frequency weighting characteristics refers.

f: the standard deviation of *n* independent repeated measurements of the A-weighted emission sound pressure level, where *n* is the number of repeated measurements as specified in the standard.

Conclusion

Within the scope of legal metrology, toy testing for safety against the international standard ISO 8124 series helps to reduce technical barriers to trade on toys among economies. Laboratory accreditation based on ISO/IEC 17025, being a common platform for showing technical competency of testing laboratory, provides an additional means towards removal of trade barriers. Apart from being a mandatory requirement of ISO/IEC 17025 for assuring reliability of laboratory testing, measurement uncertainty is vital for decision-making regarding product compliance. The ISO GUM sets out a useful and practical guideline for the evaluation of measurement uncertainty. The Hong Kong Government Laboratory has made use of GUM to set up models for estimation of uncertainty of tests prescribed in ISO 8124. Selected examples showing our efforts on the matter are presented in this paper.

Reference

1. ISO/IEC 17025:1999 *General requirements for the competence of testing and calibration laboratories*, International Organization for Standardizatio n / International Electrotechnical Commission, 1999.

2. *Guide to the expression of uncertainty in measurement*, International Organization for Standardization, Geneva, 1995.

3. IEC 60804: 1985 Integrating-average sound level meters, International Electrotechnical Commission, 1985.

4. IEC 60651:1979 Sound level meter, International Electrotechnical Commission, 1979.

Traceability of Analytical Measurement in Chemistry and CRMs-related Activities in China

Yu Yadong

National Research Center for Certified Reference Materials

No. 18 of Beisanhuan Donglu, Beijing 100013, People's Republic of China

Abstract:

In order to make the analytical measurements fit the purposes, the validity and reliability of the results is necessary. The way to obtain the validity is to accomplish the traceability. Base on this point, the author has tried to theoretically probe into the issue of traceability of analytical measurements in chemistry in this article. A model of proposed traceability system is also put into discussion. As the key element in the traceability system for chemical measurement, the role and position of CRMs are discussed, and the administration and development of CRMs in China are also introduced.

1. Introduction

It is well known that analytical measurement results are becoming more and more important as the basis for making the important decisions nowadays. Neither top-ranked government officials, who are responsible to draft the regulations and the policies, nor ordinary people dealing with the routine activities can avoid of using analytical measurement results. For example, to implement the policy for sustainable development and to pay attention to the global green house effect, government officials need to adopt the analytical measurement results so as to draw a clear picture of the environment situation. Without the analytical results, it is impossible to evaluate the matters of food safety and health care, which the public concerns the most. Analytical measurements are involved in all aspects of development of science and technology, economy and society and a very small error may result in a wrong decision, which will lead to some negative influence in a large scale and in a long term. Therefore, the reliability of an analytical measurement result is extremely important.

There exist enormous demands for the analytical measurements in the fields of scientific and technical, economic, and social development. The resources of analytical science and technology, which are the exiguous resources, can only meet a part of the demands.

Unfortunately, these exiguous resources are not efficiently utilized though some measures for quality control have been taken. A large portion of analytical measurements (say twenty to thirty percent) cannot "fit for the purposes". China, on one hand, is a big developing country and the scientific and technical resources for analytical measurements are badly short. On the other hand, non-efficiency utilization of the resources indeed exists since measurement results from different laboratories are lack of comparability and cannot be sheared among the users. The validity of the analytical measurement results needs to be improved earnestly.

Furthermore, the economy globalization is becoming an irreversible trend of world development, which has brought or will bring deep influence to all aspects of economy, science and technology, culture and social life in countries. To promote the international trade, to reduce the TBT, to protect the environment, to improve the welfare of people from various countries around the world, the multilateral adoption of analytical measurement results is required. Valid analytical measurements are increasingly becoming the technical basis for the communication and the activities related to the international trade, commerce and regulatory affairs.

To sum up, it is significant to improve the quality and the validity of analytical measurement results so as to fully meet the requirements of the measurement purposes. To accomplish the validity of analytical measurements, one of the main measures is to make the analytical measurements traceable to SI unit or accepted units in the fields of, no matter which, industrial or legal metrology.

2. The Role of CRMs in Traceability

2.1The Principles for Obtaining a Valid Result:

National Research Center for Certified Reference Materials of China (NRCCRM) has applied itself to the improvement of analytical measurements since its foundation. Based on the experiences built up for decades, the principles for obtaining a valid measurement result are brought forward as follows and illustrated in figure 1:

- A valid result should be reliable;
- The reliability of a result is from comparability;
- The comparability requests traceability.

As illustrated in the scheme above, if being valid, an analytical measurement result should be just in a position to serve the purpose of the analytical measurement. To be a valid result, it shall be reliable. However, how can a result be reliable? There is a saying, "no comparison, no judgment". So, whether or not a measurement result is reliable can only be said after

compared with other results. In the other words, the result has to have the comparability with others or else it cannot be a reliable result. The reliability comes from the comparability of the analytical measurement result. If the comparability of the analytical measurement results is required, the analytical measurements that produce these results must be traceable to some equivalent references (measurement standards, normally reference materials for analytical measurements). To make it short, a valid result of the analytical measurement has to have the property of traceability.

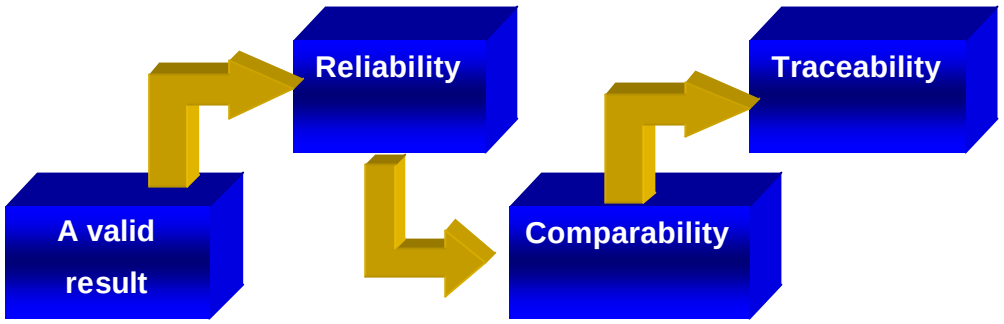


Fig 1. The scheme of the principles for obtaining a valid result

2.2Definition of Traceability:

"Property of the result of a measurement or the value of a standard whereby it can be related to stated references, usually national or international standards, through an unbroken chain of comparisons all having stated uncertainties."

Notes:

- The concept is often expressed by the adjective traceable.
- The unbroken chain of comparisons is called a traceability chain.

Between the lines of the definition, we may find out several points which are very important to understand.

- First of all, the traceability is a given property to a measurement result by the operator of this analytical measurement. The purpose for making the measurement traceable is to ensure the validity of the result. A valid result has to have the property.

Secondly, an uncertainty shall be considered into a part of the result of the analytical measurement. A traceable result should be accompanied with a stated uncertainty, or else it is not considered as integrated.

- Next, in order to relate the measurement result to some references, there should be some comparison modes that form the basic links in the comparison chain.

- Fourthly, the chain of comparisons that leads the result of a measurement finally to be related to a stated reference, normally national or international measurement standard shall be unbroken, or else the traceability of the result can not be achieved.
- At last but not at the least, ***to achieve the traceability in a country or an economic zone or a region, there must be a technical platform which is made of a group of measurement standards, including certified reference materials especially for analytical measurements in chemistry.*** As far as the traceability of analytical measurements in chemistry concerned, certified reference materials play an extremely important role as measurement standards.

3. Traceability system for analytical measurement in chemistry

3.1 A Sketch Map of Theoretical Traceability System

Herein below is the sketch map (Figure 2) of theoretical traceability system for analytical measurements in chemistry drawn by ISO/REMCO.

Fig. 2 A Sketch Map of Theoretical Traceability System

There are two columns in the figure 2. On the up side of the arrow of traceability, there is the column of laboratories. On the other side is the column of reference materials. They are the key elements in the system. In the traceability system, the reference materials and the laboratories at different levels construct the hierarchy of the system. Traceability of chemical measurements is carried through between the laboratories at different levels and reference materials are the backbone of the transition of chemical quantity values. We can see from Figure 2 that reference materials hold a leading position in the traceability process of analytical measurements in chemistry and that use of reference materials is an essential mean for realizing transition and traceability of the quantity values and for ensuring the accuracy and consistency of the analytical measurements.

From the map, some cognition can also be educed.

- None but the result vertically compared (traceable) through the unbroken chain of comparisons may have the comparability horizontally.
- Traceability of an analytical measurement result produces comparability, comparability forms the basis of reliability and that leads to the validity of the result. In one word, traceability results in validity. These are the basis of the principles for obtaining the valid results of analytical measurements.
- To realize the traceability of the analytical measurement in chemistry, a system with all necessary systematic elements is required. The system is called “Traceability System”.

3.2 Essential Technical Elements for a Traceability System in Practice

a. Quantities and units in the measurements of chemical components

The basic quantity and units in SI unit is the top end of measurement traceability. Due to the technology limitation, the unit of the amount of substance in SI units has not yet been realized into a practical level. However, the quantitative relations among this special quantity and the unit for chemical measurements and the other basic quantities and units can be simply established. The basic quantities and derived quantities, which are often involved in chemical measurements, such as Amount of Substance (n), Mass(m) and Volume(V), and the their units and ratios are listed in Table 1.

Table 1. Some SI Units and Their Ratios

Quantity	***Amount of substance***	***Volume (derived)***	***Mass***
Symbol of quantity	*n*	*V*	*m*
SI unit	mole	cubic meter	kilogram
Symbol of SI unit	mol	m^3	kg
	mol/(unit of quantity)	m^3/(unit of quantity)	kg/(unit of quantity)
(unit of quantity)/mol	mol/mol	M^3/mol	kg/mol
(unit of quantity)/m^3	mol/m^3	m^3/m^3	kg/m^3
(unit of quantity)/kg	mol/kg	m^3/kg	kg/kg

Some ratios among those above-mentioned in Table 1 are in common use for chemical measurements. The quantities, symbols, dimensions and definitions expressed by those ratios are listed in Table 2.

Table 2. Quantities in Common Use for Chemical Measurements

Quantity	Symbol	Dimension	Definition
Mole fraction	X	1	$x_B = \frac{n_B}{\sum n_i}$
Mass fraction	W	1	$w_B = \frac{m_B}{\sum m_i}$
Mass concentration	R	kg/m^3	$r_B = \frac{m_B}{V}$
Mole concentration	C	mol/m^3	$C_B = \frac{n_B}{V}$
Mass mole concentration	m_b	mol/kg	$b_B = \frac{b_B}{m_A}$

These quantities expressed by the ratios of SI units make the traceability of analytical measurements in chemistry to SI unit to be possible. By means of measuring other quantities, the measurements of amount of substances can be accomplished. This kind of measurements can be made to be traceable to the other basic SI units which are well realized.

b. Measurement Methods and Comparison Methods

Measurement and comparison methods are the relevant means to relate the various measurement standards at different hierarchies with each other. According to the principles and metrological characteristics, analytical measurement methods are clarified as follows:

- **Primary Method of Measurement(PMM):** A primary method of measurement is a method having the highest meteorological qualities, whose operation can be completely described and understood, for which a complete uncertainty statement can be written down in terms of SI units, and whose results are, therefore, accepted without reference to a standard of the quantity being measured. Its accuracy is always better than those of any other measurement methods.

Before the year of 1995 when Consultant Committee of Amount of Substance (CCQM) made the definition, the primary measurement methods had been called as definitive measurement methods.

For the measurements of amount of substance, there are five primary measurement methods recognized:

——isotope dilution with mass spectrometry

——coulometry

——gravimetry [(a) gas mixtures and (b) gravimetric analysis]

——titrimetry

——freezing-point depression determination

- **Reference Method of Measurement (RMM):** This kind of methods are those which have been systematically and thoroughly investigated, and in which exact and clear descriptions of the necessary conditions and procedures are given for the accurate determination of one or more property values, such as chemical component values. The documented accuracy and precision of the methods are commensurate with the methods used for assessing the accuracy of other methods and for certifying reference materials.
- **Validated Method of Measurement (VMM):** The analytical measurement methods which have been validated by systematic laboratory studies, and whose technical performance characteristics meet the requirements and specifications relating to the intended use of the

analytical results, such as the use of certifying working reference materials (second class CRMs). The performance characteristics which have been examined during the laboratory study are selectivity and specificity, range and linearity, limit of detection and quantification, bias and precision, ruggedness etc. These parameters should be clearly stated in the documentation of the measurement methods.

- **Routine Measurement Methods:** The operational procedures for obtaining various data are considered as this sort of methods. According to the requirements of the measurement purposes, they can be the reference methods of measurement, valid methods of measurement or other laboratory self-developed measurement methods. Usually, routine measurements are the headstreams of the measurement traceability demands.

In general, comparison methods are prescribed in the rules of verification or the calibration specifications. Comparison methods are substantially also measurement methods. According as the theoretical principles and the level of stated uncertainty, comparison methods may be clarified into High Precise Comparison Methods and Precise Comparison Methods.

c. (Primary) Measurement Standard for Analytical Measurements in Chemistry

National (primary) measurement standards are the objectives of measurement traceability and the key element of the national traceability system. Due to the particularity of the analytical measurement in chemistry, measurement standards differ from those for physical measurements. The measurement standards in the traceability chain of chemical measurements are shown in the forms of CRMs. The property value of the measurement standard is fixed and stored in the CRMs once the reference material is certified. The property value is transferred while the CRM is moved at the scales of time and space. The transition of the property value and the behavior of traceability occur when the CRM is used for the purposes of calibrating a measurement apparatus, assessing a measurement methods and assigning a value to another material.

Reference materials are divided into three hierarchies:

- **Primary Reference Material (PRM):** A primary RM is one having the highest metrological qualities and whose property value(s) is determined by means of a primary method of measurement (PMM).

In China, this sort of reference materials, which conforms with the definition of a national measurement standard in related document, is generally included in national first class CRMs. Users may make their own choices according to the requirements of analytical measurements they are going to deal with.

- **Certified Reference Material (CRM):** Reference Material, accompanied by a certificate, one or more of whose property values are certified by a procedure which establishes traceability to an accurate realization of the unit in which the property values are expressed, and for which each certified value is accompanied by an uncertainty at a stated level of confidence.

CRMs conform with the definition of national measurement standard. It is also called first class certified reference materials in China.

- **Working Reference Material (WRM)**: Materials or substances one or more property values are sufficiently homogeneous and well established to be routinely used in field of analytical measurements for calibrating an analytical apparatus, for quality control of analytical measurements, for ensuring the traceability of analytical measurements, for assessing an analytical measurement method, or for assigning values to a substance analyzed.

It is also called second class certified reference materials in China and is the measurement standard, which needs to be compared with national measurement standard, implied in the definition of traceability.

d. Measurement Uncertainty

Measurement Uncertainty: Parameter, associated with the result of measurement, that characterizes the dispersion of values that could reasonably be attributed to the measured.

From the technical point of view, uncertainty is a part of a measurement result. Without well estimated uncertainty, a measurement result is not integrated and meaningless in practice. If any comparison step is not accompanied with reasonably estimated uncertainty, the traceability chain defined in VIM will be impossible, nor will be able to exist. The information about uncertainty for a measurement standard shall be available, otherwise the uncertainty estimation for the measurement result related to the standard can not be accomplished and the traceability of the measurement can not be achieved. In despite of the difficulty to show it in a sketch map, measurement uncertainty is assuredly one of the important technical elements in the traceability system. As same as any measurement result, the production process and procedure of uncertainty estimation should be described in detail in the measurement methods and comparison methods. Owning to the fact of particularity and complexity, uncertainty estimation for analytical measurements in chemistry is very difficult. Therefore, as the key content of the study on the traceability system in chemistry, more attention must be drawn to the principle and procedure for estimating uncertainty of an analytical measurement in chemistry.

With all these essential elements, a framework of a practical traceability system for analytical measurement in chemistry can be constructed.

3.3 A Model of proposed traceability system for analytical measurement in chemistry

Figure 3 shows a model of proposed traceability system for analytical measurements in chemistry brought forward by the author. At the bottom is the metrology customers engaged in various chemical measurements and the top is SI unit, which is the top end of the traceability system for analytical measurement in chemistry. The middle arrange of the model is RMs, the standards of the measurement on different hierarchies. Both left and right sides of the model are the traceability means, the left is formed by various kinds of comparison methods and the right side is formed by different measurement methods.

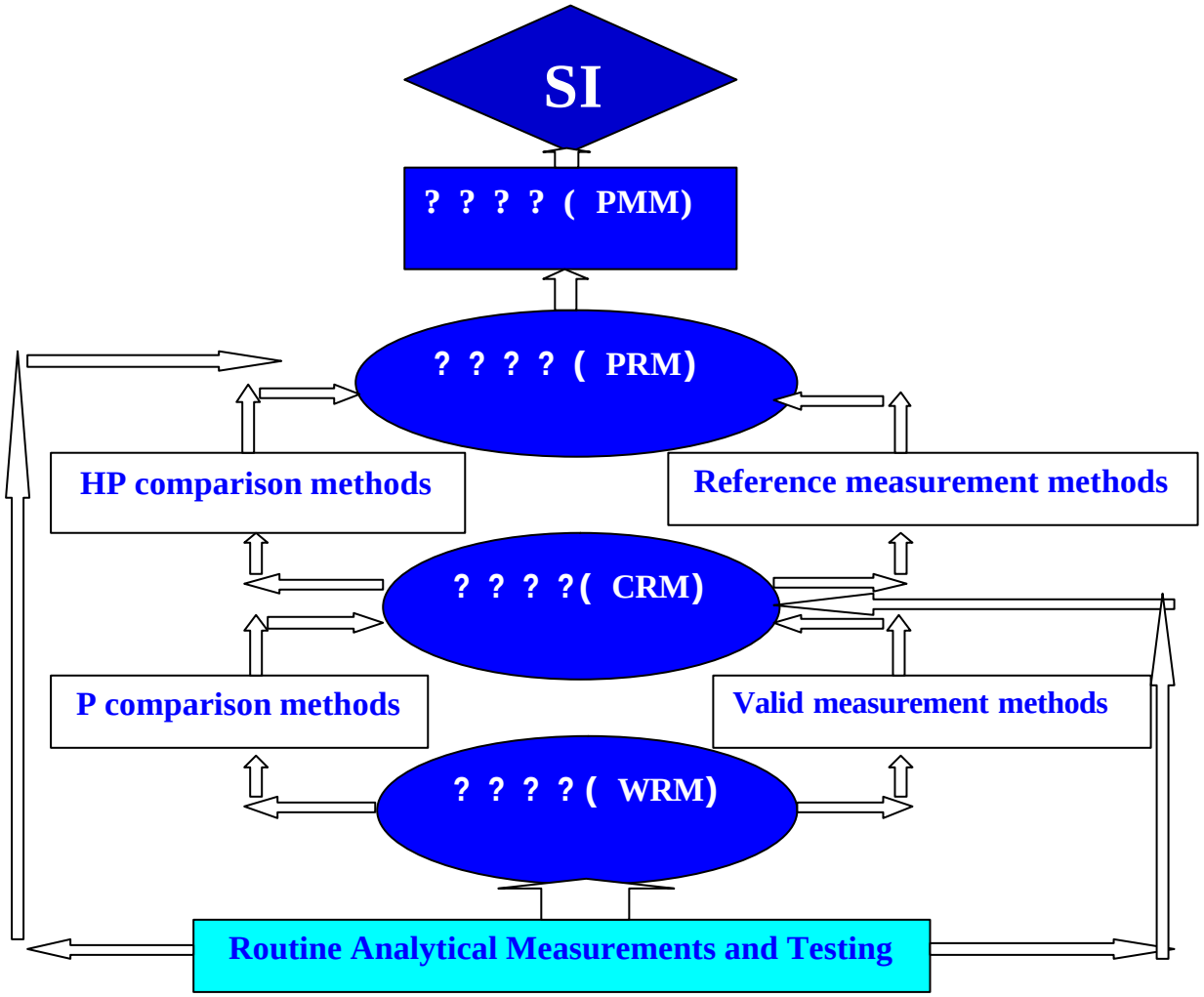


Fig.3 A Model of proposed traceability system

From this model, it can be easily seen that the implementation of the traceability for analytical measurement in chemistry seems like a climbing, from the bottom to the top, a pyramid. The reference materials, which are the measurement standards act as the daks or stations on the way of climbing. At the same time, different methods of comparison and measurement work as the bridges between the daks. In this way, the climber of this traceability pyramid could select their routes according to their own needs and finally got to the top of it, actualize the measurement traceability.

3.4 The implied dynamic mechanism of traceability system

Figure 4 is a sketch of the dynamic mechanism of traceability system brought forward by the author.

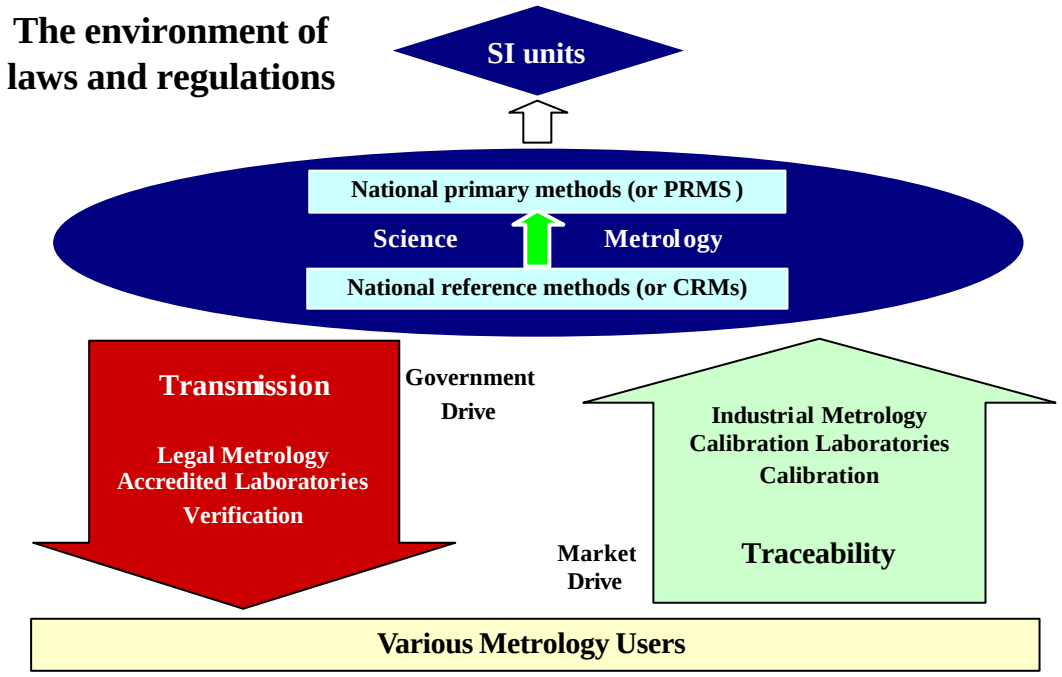


Fig.4 The implied dynamic mechanism of traceability system

The government has the responsibility to ensure the accuracy and consistency of the relevant quantity values in that some measurement results are concerned with the fair trades, public security and other public profits. The way to do this is to hand on the quantity value of the reliable measurement standards. The executors of the transferring for the quantity values are the legal metrological bodies authorized by the governing body of the governments. They pass the quantity values to the users of the measurement instruments concerning with the fields of commerce, personal security and environment protection etc. The major means of

transferring of quantity values are verification and it is driven by the government by establishing the administrative and technical rules and the regulations. The transferring of the quantity values is subject to the category of legal metrology.

Comparatively to the transferring of quantity values is the traceability of the quantity values. It is accomplished by the measurement executor including different accredited calibration laboratories and other metrology users dealing with various measurements and bringing out the data. In order to be outstanding in the market competition, the analytical measurement makers have to ensure the quality of their data and this results in the demand of the traceability of measurement results. This need comes from the pressure of the competition in the market and influenced by the social quality culture formed in the market economy system. The realization of the traceability is driven by the market and the major way to accomplish the traceability is calibration and the comparison of the measurement methods. The executor and the beneficiary of traceability are the same because the executors are also the metrology users and the laboratories providing the calibration service. Therefore, traceability activities are more active than that of the transferring of quantity values. The traceability of measurement results is subject to the category of industrial metrology.

In order to ensure the transmission and traceability of quantity values and provide technical guarantee to legal and industrial metrology, (primary) measurement standards (including CRMs) have to be developed and established. This is also the major task of scientific metrology.

Scientific metrology is an important composing part of the national basic scientific work. It involves the top part of the quantity transmission system and the traceability system, and represents the highest national measurement capacity. Many industrial laboratories and special enterprises have taken part in the scientific metrological activities in China, except the national metrology laboratory, so that a dispersed (primary) measurement standard system has been developed especially in the field of analytical measurement in chemistry.

The direction of transmission of quantity and traceability is opposite and driving force is different. Under the system of the planned economy, the government undertook the whole thing so the transmission of the quantity values has been emphasized excessively. As the establishing and perfecting of the market economic system, the demand for metrology and traceability becomes stronger and stronger, especially after China accession into WTO. After the opening up of the technical market of calibration and verification, there must appear a complexion of competition and development, this will actively accelerate the improving of the capability of national measurements.

4. The Administration of CRMs in China

4.1 The Categories of CRMs

ISO/REMCO has assorted certified reference materials into 17 categories, but in China, they are divided them into 13 categories in accordance with national managing requirements. The some reference information, such as categories, serial number and amount of national certified reference materials, are listed in table 3.

Table 3 The Schedule of Categories, Serial Number and Amount of National Certified Reference Materials

Serial No.	Category	CRM (first class)	CRM (second class)
GBW 01101~ 01999	Ferrous Metals	180	147
GBW 02101~ 02999	Non-Ferrous Metals	175	10
GBW 03101~ 03999	Building Materials	67	2
GBW 04101~ 04999	Nuclear and Radioactivity	82	31
GBW 05101~ 05999	Polymer	2	3
GBW 06101~ 06999	Chemical Industry	41	321
GBW 07101~ 07999	Geology and Ores	237	165
GBW 08101~ 08999	Environmental Chemistry	229	504
GBW 09101~ 09999	Clinical Chemistry	46	17
GBW 10101~ 10999	Food Chemistry	5	8
GBW 11101~ 11999	Energy Resources	25	18
GBW 12101~ 12999	Technology and Engineering	7	22
GBW 13101~ 13999	Physico-Chemistry	70	177
Total		1166	1425

Note: The first column is the range of the serial number of CRMs, the second column lists the categories of the CRMs, and the third and forth columns list the amount of the first class and second class CRMs (by the end of year 2002).

4.2 The Classification of CRMs

China has ranked CRMs into two classes and they both accord with the definition of the CRMs. The technical requirements of the first class and second class CRMs are compared in table 4.

Table 4

Criteria / Class / Requirements	First class CRMs	Second class CRMs
Preparation Bodies	National laboratories, authoritative laboratories or other institutions	Industrial and corporations' laboratories, institutes as well as other research bodies
Certified methods	1) Use of primary measurement method, operated and certified by two analyst independently. 2) Use of two or more independent and reliable analytical measurement methods for certification.	1) Use at least one accurate and reliable analytical measurement method for certification. 2) Use of the method to compare with the first class CRMs for certification.
Measurement Accuracy	Improve the accuracy as much as possible according to the final practical requirements and the economic principles. Try to make the accuracy three to ten times as it is required in the practical requirements.	The accuracy should be three to ten times more than the accuracy of analytical measurements on site.
Homogeneity	Depend on the final practical requirements.	Depend on fthe final practical requirements.
Stability	The stability limit should beat least 1 or 2 years.	If the CRMs can be acquired and used immediately, the stability limit could be short to only several weeks.
Main uses	1) For study or evaluation of the reference methods. 2) For certification of the second class CRMs. 3) For calibration of the measurement instruments with high precision and accuracy.	1) For study or evaluation of the measurement methods. on site. 2) For quality assurance of the field laboratories. 3) For quality assurance among different laboratories.
Packaging	The packaging should follow the requirements of the technical rules of CRMs	The packaging should follow the requirements of the technical rules of the CRMs

4.3 The Coding of CRMs

The former National Metrology Bureau had established " The Coding Method for Certified Reference Materials" in the year 1984, which prescribed that we use each of the first capitalized block letter " GBW "in the Chinese pinyin of the words " National, reference and material " as the head for the code of the national reference materials. The code can be seen as follows in Figure 5:

At the beginning of a code is the code head of the reference materials " GBW ", which stands for the national reference materials. After that, letter X represents the number of the reference material in the first category with two digits. Then, there is the letter Y, which represents the number of the reference material in the secondary category with one digit, so that there are nine categories here. The letter Z is following Y and it stands for the order of the RMs after the secondary sort with two digits. The order is arranged according to the authorized time. The last letter U represents the batch number of the RMs production shown with a small English letter and the order is the same as in English.

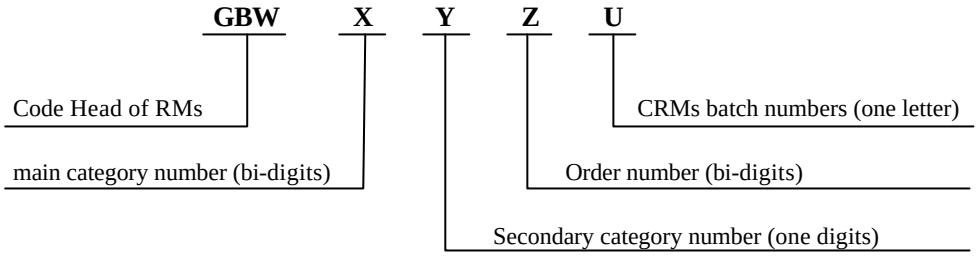


Fig 5 The coding of the first class reference materials

4.4 The Documents for CRMs Administration

The administration of the CRMs is accomplished in accordance with the law and the regulations prepared by the responsible government department. Under the law of metrology, there is two kinds of documentations. The first kind of documentation is the administrative rules that include "The means of examining and authorizing to produce GBW CRMs" and "The rules for administration for GBW CRMs". The second kind is the technical regulations that comprise "The method of numbering GBW CRMs", "The regulations of compiling certificates of GBW CRMs", "The general terms of GBW CRMs" and "The technical norm of the preparation and certification for GBW CRMs". They are listed in Table 5.

The Law of Metrology	
Administrative Rules	**Technical Regulations**
Means of examining and authorizing to produce GBW CRMs	**Method of Numbering GBW CRMs**
	Regulation of compiling certificate of GBW CRMs
Rules of Administration for GBW CRMs	**General Terms of GBW CRMs**
	Technical Norm of the preparation and certification for GBW CRMs

Table 5 The law and the regulations for the administration of CRMs

4.5 The approval procedure for CRMs

The approval procedure of CRMs is shown in Figure 6. At the bottom of the map, there is the CRM producer, whose responsibility is to develop and supply CRMs and hand in the documents of them to the administrative office in the middle. The main duties of the office is to issue CRM certificates and production license, edit catalogue of CRMs and check the documents received from the CRM producers for approval. Then, they pass the documents to NRCCRM, whose job is to carry through the primary review of the documents and organize the final review by calling in the CRM expert group meeting. If the CRM documents pass the final review, the materials will be handed back up to AQSIQ through the administrative office to get the approval. After approval by the AQSIQ, the office will issue the certificate and the production license so that the CRM can be sold in the market. CRM producers have to ensure the supply and the quality of the CRMs, and make sure that they are traceable to NRCCRM.

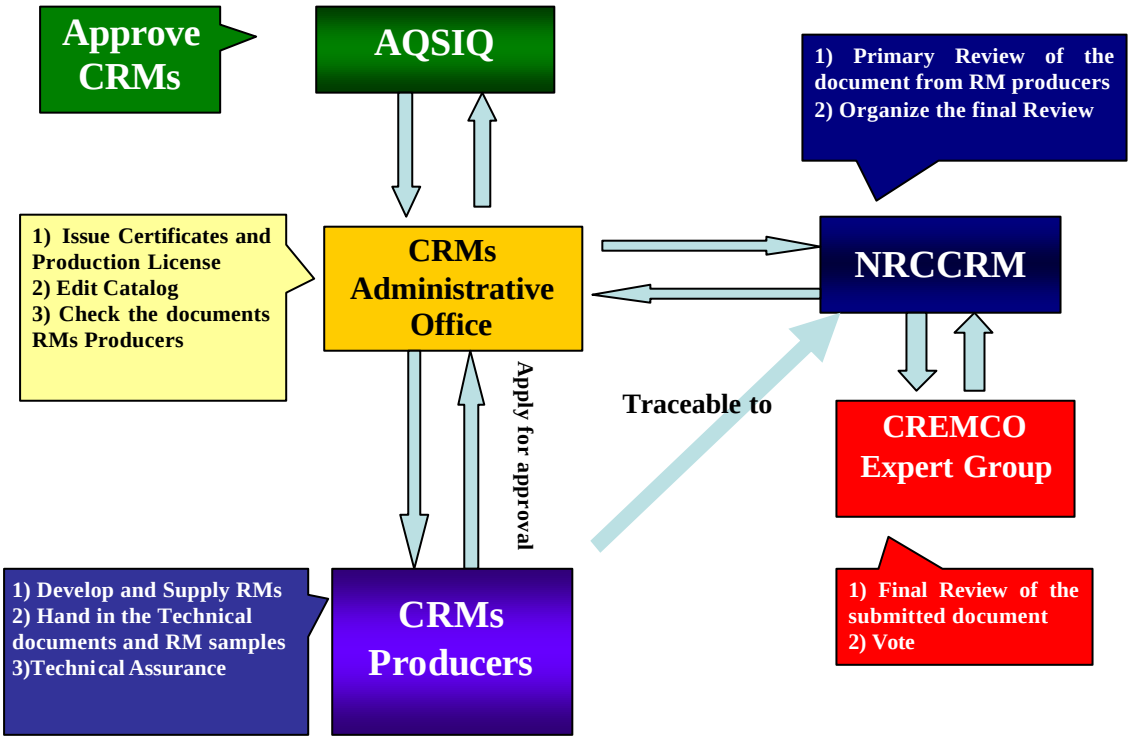


Fig. 6

5.The Development of CRMs in China

Since the first CRM was issued by Chinese government in 1951, metrological activities in chemistry have been gone through and developed over fifty years in China. Six national

primary measurement standards have been established in NRCCRM and more than 2591 kinds of GBW CRMs are available in China by the end of 2002. They are classified according to metrological characteristics into the primary reference material (PRM), first class certified reference material (CRM) and second class certified reference material (working RM). And based upon the application fields, they are assorted into 13 categories, which are showed in table 1. Fig 7 shows the development on the number of CRMs in China from 1983 to 2002. More than 100 verification regulations for different instruments for analytical measurements have been issued since 1949. Now, a national analytical measurement system is being formed in the country with Chinese characteristics.

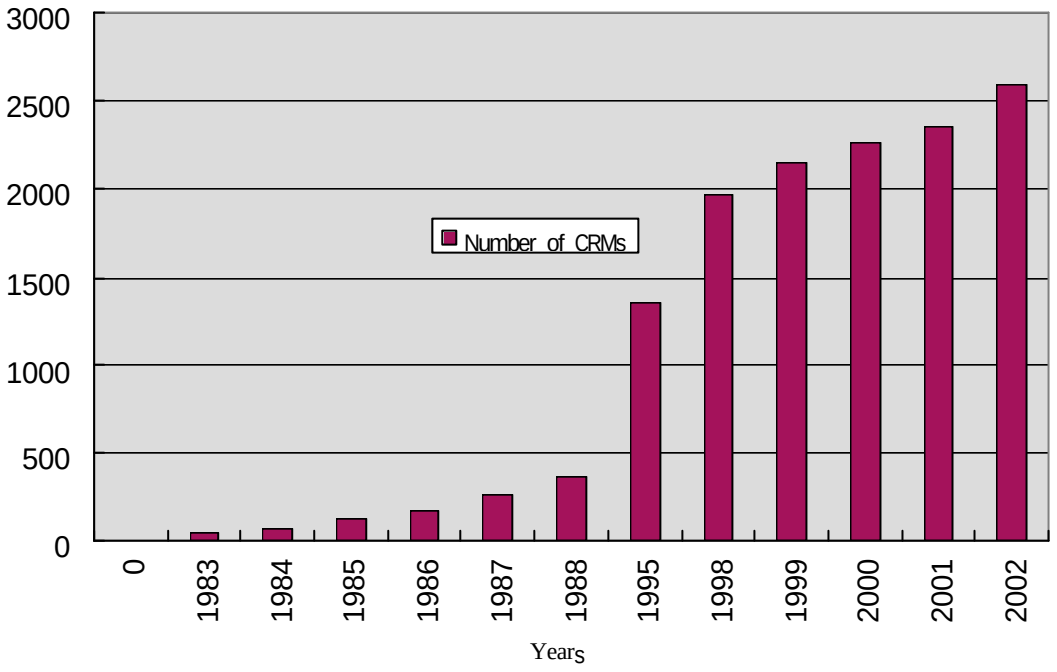


Fig.7 The Development of GBW CRMs in China

In recent years, the users of GBW CRMs have extended to about 30 countries around the world, such as USA, Canada, UK, France, Germany, Italy, Spain, Sweden, Australia, New Zealand, South Africa, Japan etc.

It is clear to see from Table 6 that the amount of RMs in China is large and in fact, till the year 2002, the number of the CRMs approved by the authority has been more than any other country in the world. But on the other hand, we have widely gap with the advanced countries in the distribution of CRMs, especially in the field of the advanced technology. In one word, China will have to make great effort in the research and development of CRMs.

Table 6 The Percentage and Distribution of CRMs in CORMA for Countries

Category / Counties	Ferrous Metals	Non-Ferrous Metals	Inorganic	Organic	Physical and Technological Characteristics	Biological and Clinical	Life Quality	Industry	The Amount in the CRM Bank
China	11.46	6.00	18.28	6.56	7.87	6.85	9.02	24.71	9.13
France	10.48	18.17	8.02	10.37	23.14	2.49	8.08	5.08	14.06
Germany	5.65	16.52	7.54	2.62	10.09	1.25	0.40	6.73	9.79
USA	9.01	5.54	10.75	15.75	9.27	16.51	4.51	28.80	10.13
UK	27.90	27.25	5.69	20.87	32.57	0.62	13.06	6.68	23.22
Japan	18.84	0.64	17.16	15.88	2.09	0.94	23.84	1.94	7.60
Russia	9.50	12.49	9.46	0.26	0.43	0	9.09	7.32	7.46
International Organizations	0.22	1.74	1.68	12.73	10.34	66.67	17.31	3.19	7.50
Slovakia	0.87	0.03	7.30	10.63	3.21	1.56	6.94	5.23	2.28
Holland	0	0	3.05	0.66	0.54	0	4.92	0.60	0.84
Canada	0.22	1.04	1.92	3.28	0.59	1.56	2.83	4.04	1.55
S. Africa	0	0.41	2.17	0	0	0	0	2.69	0.60
Other Countries	5.86	10.17	6.98	0.39	2.04	1.56	0	2.99	5.84

Note: the data in the table above is in accordance with the CORMA (97 edition).

6. The Role of NRCCRM

6.1 Basic information on NRCCRM

China National Research Center of Reference Materials (NRCCRM) is subject to the State General Administration for Quality Supervision, Inspection and Quarantine (AQSIQ). It is a technical body with its mission "to ensure the uniformity of the measurement units and the accuracy of the measurement results by studying on the analytical measurement in chemistry, so as to underpin the national development in the fields of economy, society, science & technology etc."

NRCCRM has passed the accreditation by CNAL in 2000. From very beginning as the representative of China, NRCCRM participates actively in the international or bilateral comparisons and has got good results in the most cases.

6.2 Main tasks

In the establishment and improvement of the National Measurement System in Chemistry, which forms a part of the global Measurement System, NRCCRM is assigned with three main tasks:

- First, to keep the national capability of analytical measurement in chemistry. This should be advanced and competitive internationally.
- Second, to develop and improve the national measurement standards, including Certified Reference Materials (CRMs). These standards should also have the comparability and equivalence around the world. and
- Third, to establish and maintain a traceability system for analytical measurements in chemistry, which allows the metrology users to access and use the national measurement standards.

6.3 Special achievements

During these years, NRCCRM has made quite a few achievements including:

- High accurate measurement methods have been studied. Six elements' atomic weights were determined and accepted by the international organization.
- 6 national primary standards have been established, including combustion heat, acidity, conductivity, viscidity, humidity, purity of primary standard reagent. More than 40 national standards relating to chemical compositions, physical chemistry and chemical engineering characteristics have been developed.
- More than 400 first-class and second-class national certified reference materials have also been developed.
- NRCCRM has participated many international comparisons, including IDMS method, primary gases and primary reference reagents organized by CCQM and other organizations and has got good reputation.
- Involved in many national research programs and has got more than 50 national and provincial technology advancement awards for high level achievements.

6．4 Recent Research Projects

On the basis of our former development of the work concerning the CRMs in China, NRCCRM will do its best to improve capability of analytical measurements in chemistry and measurement quality. In order to get further development, recently we are going to carry on some research projects mainly as followings:

a. Establish and improve National Analytical Measurement System (NAMS), which was supported by national ministry of technology and science.

b. Study and establish of national food safe monitoring system, especially measurement standards, which was supported by national ministry of technology and science.

c. Study of the inspection techniques of the active components of food, Chinese traditional

medicine and natural herbal medicine, which was supported by national ministry of technology and science.

d. Study of the traceability system of the pesticide inspections and measurements

7. Conclusion

To sum up, the conclusions can be drawn out from above discussion as follows:

- Traceability is a fundamental issue in metrology. CRMs are the key elements in the traceability system.
- An administrative system, both in documented form and in actual form, has been constructed in China.
- China CRMs development started late but has gone fast. However, distribution should be improved.
- NRCCRN plays a leading role for the Chinese CRMs development in the past, now and will do the same in the future.
- Several scientific projects, such as the establishment and improvement of NAMS, have been started-up and more work has to be done. From the technical point of views, international communication and cooperation is called on.

Mutual Recognition Arrangement by the CIPM for Establishing Global Traceability of Measurement - its Role and Mechanism -

Kozo Iizuka

(Past Vice-President of the CIPM)

President of Japan Association for Metrology Promotion (JAMP),

25-1 Nando-cho, Shinjuku-ku, Tokyo 162-0837, Japan

1. Introduction

The history of the measurement system goes back to the ancient civilization of several thousand years ago. However the modern measurement system on a global scale was first established by the Metre Convention (la Convention du Mètre) only about 200 years ago. The objective of the Convention was originally to unify weights and measures by the metric system but is now to ensure world-wide unification of measurements based upon the SI (International System of Units), which includes almost all kinds of physical and chemical units. In this paper, the recent development of the structure and the function of the organizations set up by the Convention together with their effect on legal metrology is reviewed. Then the process to establish the mutual recognition arrangement of national standards and calibration certificates and its role for world-wide traceability are explained. Finally, the effect and future of the arrangement is discussed.

2. Development of the Metre Convention

Fig. 1 shows the structure of the organizations set up by the Metre convention. The structure has been the same since 1927, when the first consultative committee was created.

The General Conference of Weights and Measures (CGPM) consists of the delegates from 51 member states and makes important resolutions concerning the new definitions and the improvement of the unit system, SI, and necessary arrangements for promoting the unification of measurement and for the operation of the International Bureau of Weights and Measures (BIPM).

The International Committee of Weights and Measures (CIPM) consisting of 18 experts supervises the BIPM under the authority of the CGPM and makes decisions on

scientific and practical problems concerning metrological standards. The CIPM entrusts detailed investigations to its Consultative Committees (CCs) set up for each field or issue. There are now 10 Consultative Committees as is shown in Table 1.

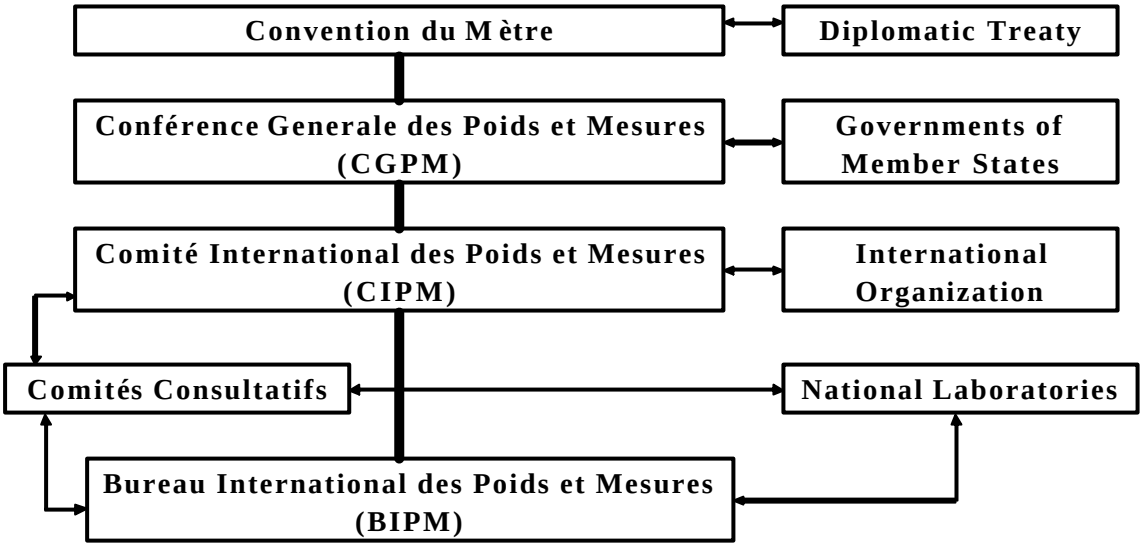


Fig. 1 Structure of the Metre Convention

Table 1 The Consultative Committees (CCs) of the CIPM (Chairs in October 2003)

	Name: Consultative Committee for ……	Acronym	Start	Chair
1	Electricity and Magnetism	CCEM	1927	B. Inglis
2	Photometry and Radiometry	CCPR	1933	F. Hengstberger
3	Thermometry	CCT	1937	H. Ugur
4	Length	CCL	1952	M. S. Chung
5	Time and Frequency	CCTF	1956	S. Leschiutta
6	Ionizing Radiation	CCRI	1958	G. Moscati
7	Units	CCU	1964	I. Mills
8	Mass and Related Quantities	CCM	1980	M. Tanaka
9	Amount of Substance	CCQM	1993	R. Kaarls
10	Acoustics, Ultrasound and Vibration	CCAUV	1998	A. J. Valdes

Corresponding to the development of science and technology, the needs for the metrological standard has expanded in wider areas and the CIPM created new CCs one by one to make thorough investigations in each field. In fact, the present names of the CCs do not fully represent their roles, as a number of working groups are attached in a limited number of CCs, even if their tasks are not exactly in the respective fields. As an example, 9 Working Groups of Consultative Committee for Mass and Related Quantities (CCM) studying on different topics are shown in Table 2.

Table 2 Working Group of the CCM in 2003

Name of Working Group	Chair
Mass standards	M. Gläser(PTB)
Density	K. Fujii(NMIJ)
Force	M. Peters(PTB)
High Pressure (above 1Mpa)	J.-C. Legras(BNM-LNE)
Medium Pressure (1kPa-1Mpa)	P. Leggat(NPL)
Low Pressure (1Pa-1kPa)	A. P. Miller(NIST)
Avogadro Constant	P. Becker(PTB)
Hardness	A. Germak(IMGC)
Fluid Flow	G.E. Mattingly(NIST)

Further, some working groups have sub-groups or task forces under them. Thus, the field covered by one CC is much more extensive than their names. For instance, Fluid Flow Working Group in Table 2 has 6 sub-groups for different kinds of flowmetry as shown in Fig. 2 and is working like an independent consultative committee. The biggest and developing CC is Consultative Committee for Amount of Substance (CCQM). It was originally created to deal with chemical standard but now has expanded its activity in the measurement of biological quantities by cooperating with other international organizations such as World Health Organization (WHO) and International Federation of Clinical Chemistry and Laboratory Medicine (IFCC). The working groups under this consultative committee are shown in Table 3.

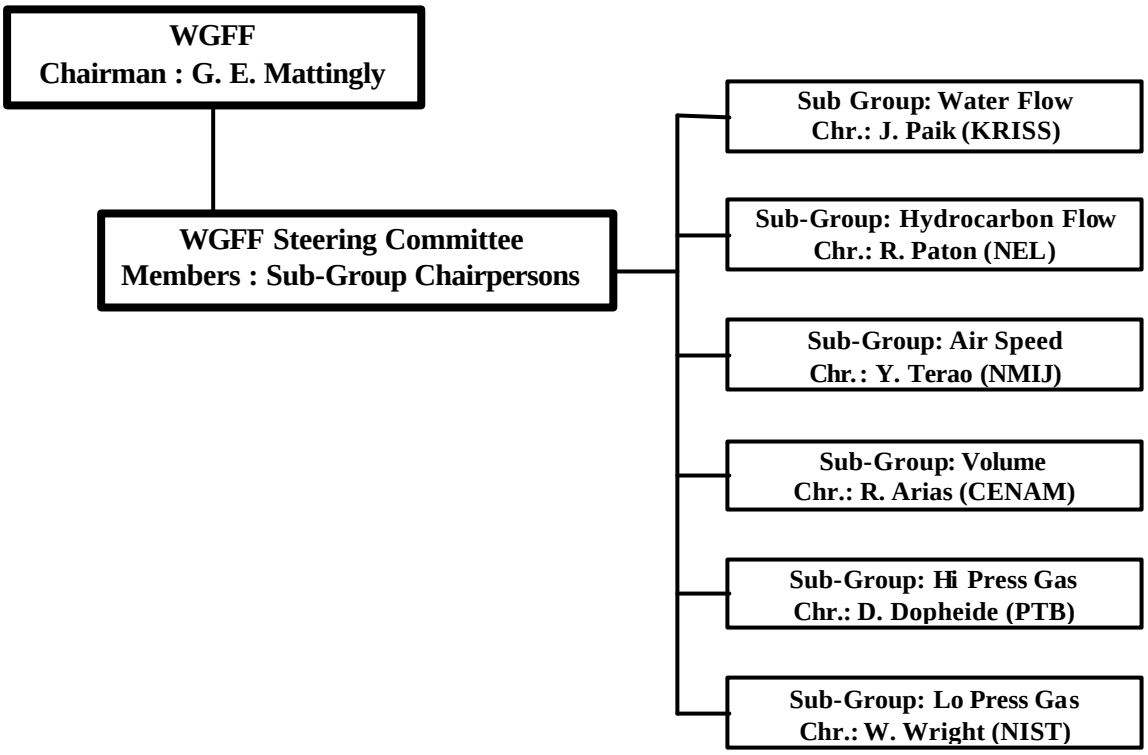


Fig. 2 Working Group for Fluid Flow

(as of October, 2003)

Table 3 Working Groups of the CCQM in October 2003

Name of Working Group	Chair
Key Comparis on & CMC Quality	J. McLaren (NRC)
Gas Analysis	E. de Leer (NMi)
Inorganic Analysis	M. Sargent (LGC)
Organic Analysis	W. May(NIST)
Electrochemical Analysis	M. Mariassy (SMU)
Bio-Analysis	H. Parks (LGC)
	V. Vilker (NIST)
(Ad Hoc)	
Surface Analysis	M. Seah (NPL)

Traditionally, the CIPM has concentrated its activity in the problems of primary standards, especially of the base quantities, and scientific metrology for many years. However, the requests from other international organizations related to the accreditation for the assurance of the equivalence of measurement encouraged the CIPM to consider the traceability of practical measurement. Some national metrology institutes (NMIs) proposed the creation of new CCs to investigate the standards for such practical metrology as the measurement of flow, viscosity, hardness, vibration, ultrasonic power and so on. As is shown above already, the present activity of the CIPM and CCs more or less covers all the scientific and technical areas where the standards and the traceability of the measurement are required. At the result, the CIPM is now much more involved in the calibration of the working level standards than the past.

3. Legal metrology and calibrations

In the regime of legal metrology, every decision must be stable and uniform everywhere. Otherwise the credibility for the legal authority will be lost. In the case of the verification of instruments, exact and stable metrological standards traceable to the national primary standard must be used. The technical guide on the traceability chain for legal metrology is documented in OIML/D5. The term "traceability" is defined as "property of the result of a measurement or the value of a standard whereby it can be related to stated references, usually national or international standards, through an unbroken chain of comparisons all having stated uncertainties". The definition is under the review by the Joint Committee on Guide for Metrology but may not be changed drastically. In any case, verification standards must be calibrated according to the procedure specified in the standard or recommendations for each kind of instruments. In that sense, measurements by the verified instruments are said to be traceable to the national or international standards.

4. Mutual recognition of measurement

If we refer to the traceability chain shown in the document OIML/D5, the international compatibility of the measuring instruments verified by a legal procedure can be realized only when the national standards are equivalent each other. Sometimes, working standards and verification standards could be calibrated in different countries. Then, the compatibility of these practical standards must be assured through calibrations with reference to the same or an equivalent (primary or secondary) standard. The requirement for the equivalence must be checked by making intercomparison of the national primary and/or secondary standards.

Since the accreditation of quality and testing became a popular process, the equivalence of measurement standards has been one of the key issue to guarantee the credibility of the accreditation. Thus, the CIPM and the BIPM were asked to take appropriate action to realize the equivalence of measurement standards on the global aspect. Table 4 shows the chronology of the development of the quality system and the movement of the CIPM towards the establishment of the world-wide traceability of measurement.

Table 4 Chronology of Conformity Assessment, Standards and Traceability

Year	ISO Standards on Management System	Organization Concerned with Conformity	Activity related to Traceability of Measurement
1975		ISO Guide 25	
1977		1st meeting of ILAC	
1979	Creation of ISO/TC176		
1985		Start of ISO/CASCO	
1987	ISO 9000s		
1990			Start of CIPM/WG on Chemical Measurement
1993	Creation of ISO/TC207	Start of IAF	Guide on Uncertainty
1994	Revision of ISO 9000s		
1995		Start of WTO	CGPM Resolution on Equivalence of Standards; 1st Meeting of CCQM
1996	ISO 14001	Incorporation of ILAC	
1997			1st Meeting of NMI Directors
1998			Provisional Signature of MRA
1999		ISO 17025	Signature of MRA

In 1995, the CGPM decided to investigate a measure to materialize the international traceability and set up a working group for it. It was discussed at the NMI-directors' meeting first held in 1997 and an idea to establish a mutual recognition agreement among national metrology institutes was formulated in 1998. The draft was repeatedly discussed and modified by the CIPM members and NMI directors and finally, the draft was approved at the 21st CGPM in 1999. The NMI directors representing 39 member states and the delegates

from 2 international organizations signed the agreement on this occasion (the name was changed as "arrangement" to mitigate compulsory nature). From the beginning of the drafting, it was a prevailing idea to include in the recognition arrangement not only the equivalence of national measurement standards but also the equivalence of the certificates of calibration and measurement issued by national metrology institutes. The signatory of the MRA reached to 56 states, economies and international organizations now.

The text of the MRA comprises Main Articles, Technical Supplement and 5 Appendices as shown in Table 5. As the governing and coordinating body for the practical operation of the MRA, regional metrology organizations (RMOs) and the BIPM constitute a joint committee (JCRB) and keep up-dated data in the appendices, although the CIPM bears the final responsibility for the decisions.

Table 5 Contents of the Mutual Recognition Arrangement

Main articles	Preamble and 15 articles
Technical Supplement to the arrangement	10 items
Appendix A	List of signatory NMIs with logos
Appendix B: B1 B2 B3	Results of CIPM key comparisons Results of RMO key comparisons Results of supplementary comparisons
Appendix C (Recognized certificates of participating institutes)	Quantities, ranges and calibration & measurement capability of institutes
Appendiz D	List of key comparisons
Appendiz E	Terms of reference of JCRB

The equivalence of national standards must be confirmed by international key comparisons organized by CCs and RMOs, which are illustrated in Fig. 3. The procedure of the CC key comparisons is laid down in the Guidelines for the Key Comparisons formulated by the Director of the BIPM with the cooperation of CC members. It is an elaborate process as is shown in Fig. 4, because the evaluation of the uncertainty of high level standards can be made only by comprehensive and deliberate international comparisons. On the other hand, the recognition of the calibration and measurement certificates is registered mainly by the examination of the results of RMO comparisons by the JCRB according to the procedure shown in Fig. 5.

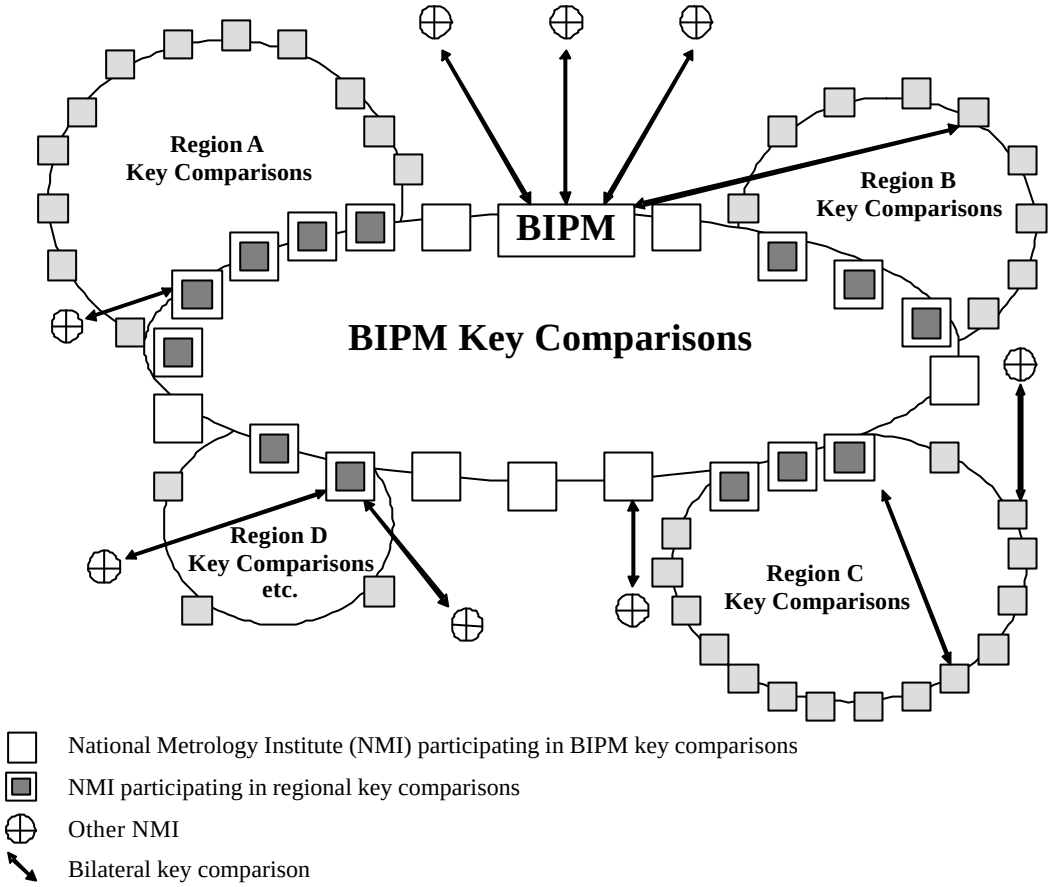


Fig. 3 BIPM, Regional and Supplementary Key Comparisons

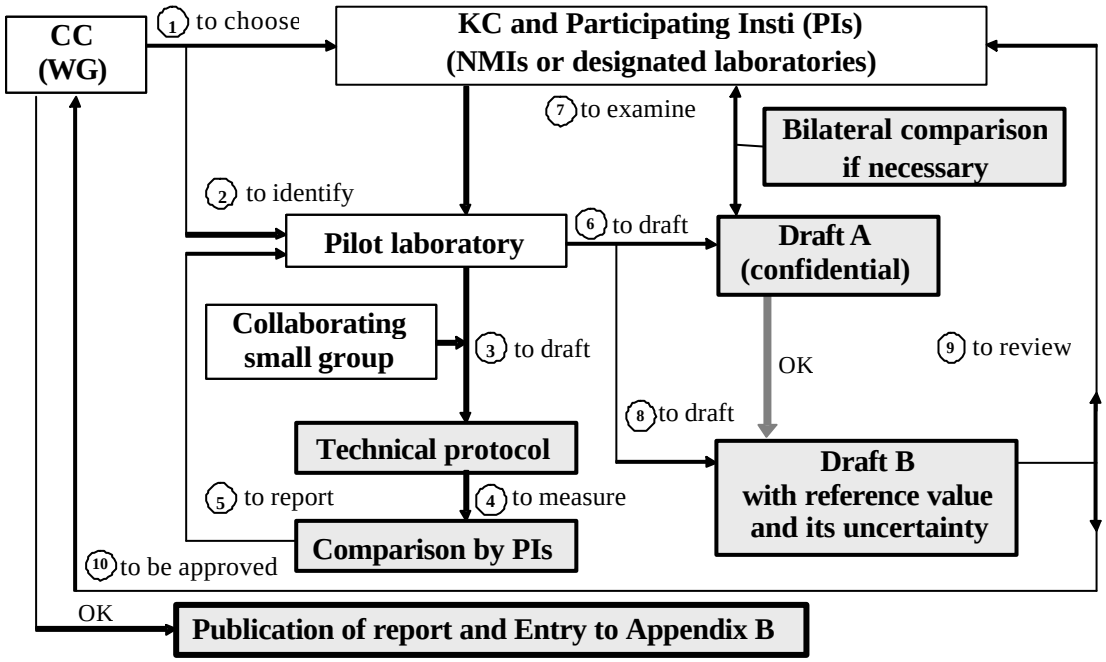


Fig. 4 Procedures for Key Comparisons

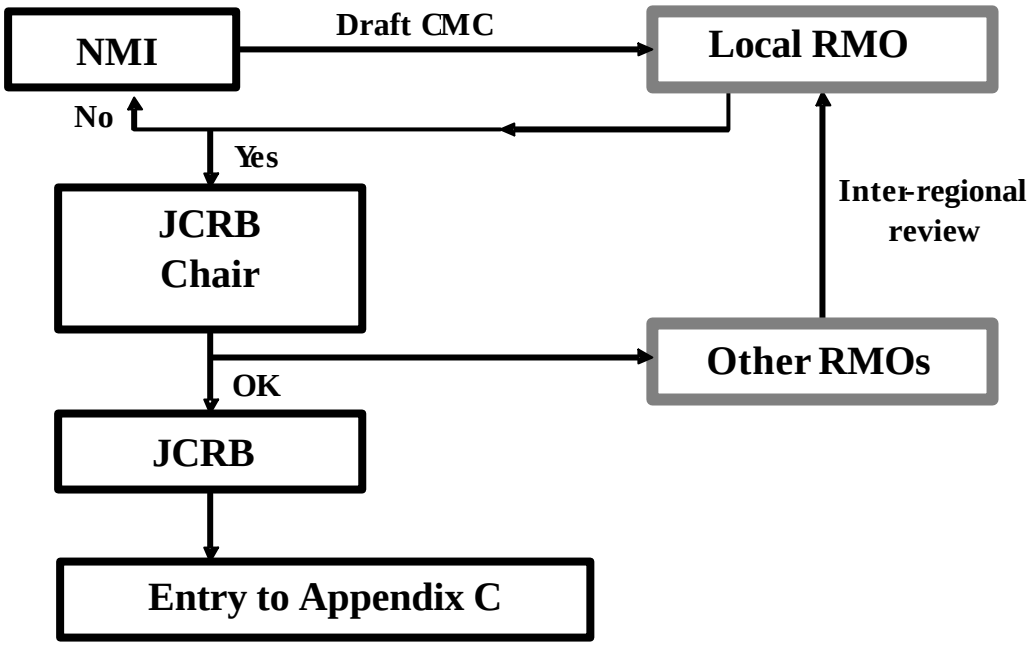


Fig. 5 Procedure for CMC Entry to Appendix C

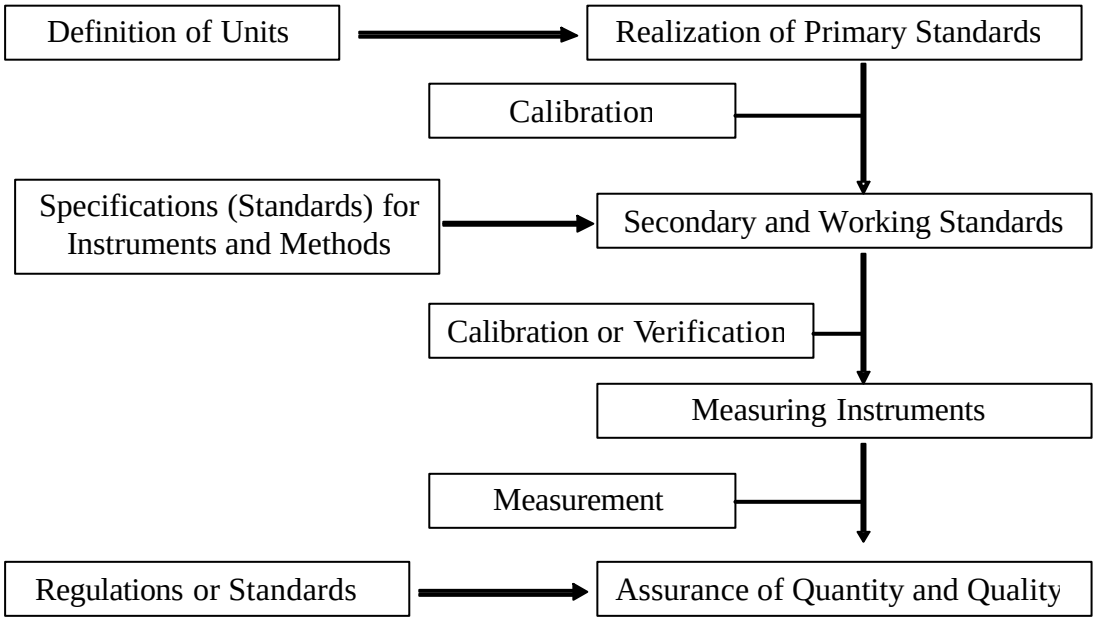


Fig. 6 Practical Implication of Hierarchy of Measurement Standards

The MRA explained above is only concerned with the standards of NMIs and calibration and measurement made by metrology institutes designated by respective national authority. In the field of legal metrology, a similar scheme is employed to ensure the equivalence of the verification of measuring instruments, because it is a part of the hierarchy of measurement shown in Fig. 6. The equivalence of verification can be guaranteed only when the verification standards are equivalent. Of course, the equivalence is guaranteed within the degree of the uncertainty of the verification standard, resulting from its calibration. Usually uncertainties of the indications of standards are negligibly small compared to the tolerance for the verification but one should always take care of the ratio of the uncertainty of the standard to the tolerance for the verification in view of the propagation of uncertainties in the hierarchy of standards.

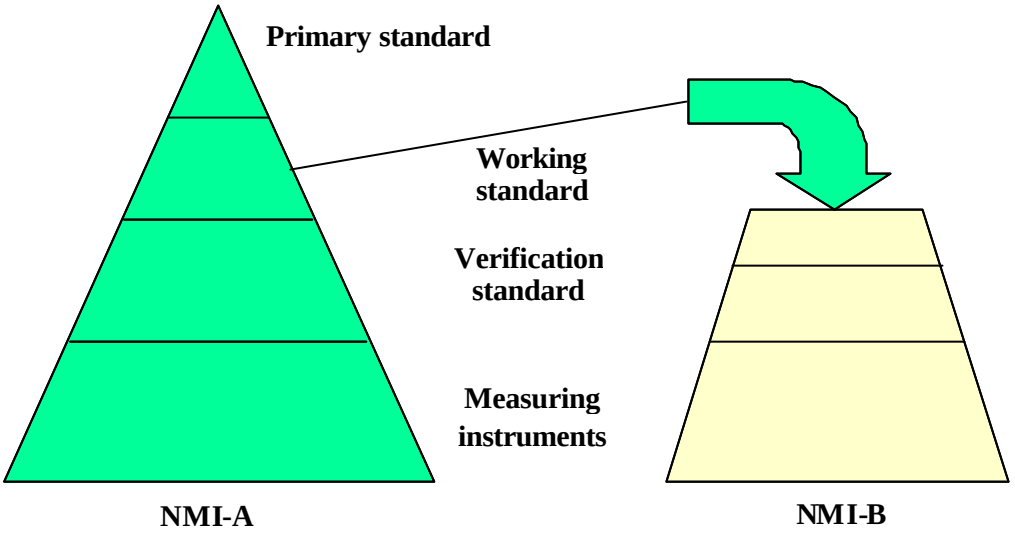
5. Present and future problems of the MRA and conclusions

(1) When the MRA went into effect in 1999, the period of 5 years up to 2003 was considered as a transitional or provisional stage. During the time, the database of the MRA, especially Appendices has been well developed by the BIPM and made available on the web (http://www.bipm.fr). The data in Appendix B and C are the products of the key comparisons organized by CCs and RMOs and those of the evaluation by the JCRB. The database is growing and is revised day by day. For the full-fledged phase, necessary key comparisons must be finished as scheduled or even replicated. At the same time, comparisons of standards in such new fields as pollution control of environment, laboratory medicine, food analysis etc. are being tackled. For establishing the traceability of chemical measurement for medicine, a joint committee with IFCC and WHO was created already. It is expected that a greater number of institutes, which are new in the field of metrology, will be included in the list of the MRA in the near future.

(2) In principle, the evaluation of the calibration and measurement capability (CMC) of participating institutes must be made on a common basis. However, the traceability system may not be the same for all states or economies. As is shown in Fig. 7, the word “working standard”, for instance, may be used for those standards with different level of accuracies as the result of different calibration hierarchy. In that sense, the terminology must be deliberately standardized. Furthermore, the realistic procedure to maintain standards and to perform calibrations must be uniform everywhere. The application of the quality system, especially ISO 17025 and related documents, may be useful and effective for solving this

problem. In fact, calibration and measurement, especially at higher accuracies, require well-trained operators. Documents help us greatly but not perfectly. In some cases, there could be inconsistency in the evaluation of the equivalence, because of a variety of the technical level of institutes and the skill of their staff. It must be a longstanding problem for the assessor of technology and a clear guideline must be prepared and followed with the consensus among concerned experts.

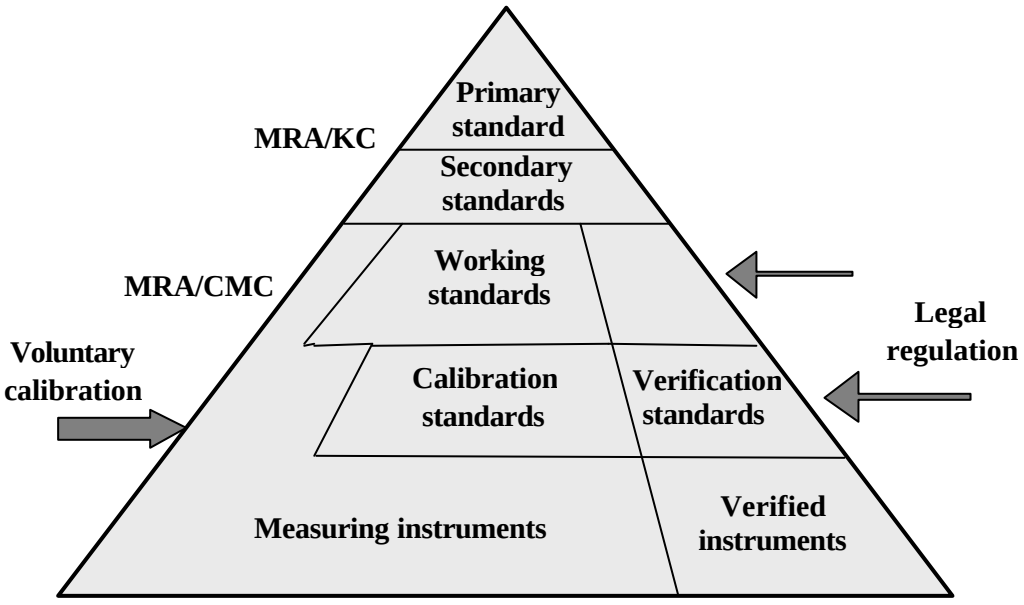
Fig. 7 Different hierarchy of calibration chain



(3) Another problem of the MRA is whether it can be a real global scheme or not. As is well known, the signatory states of Metre Convention is only 51 now, less than the members of the OIML (60 members and 49 corresponding members) and much less than those of the ISO (146 member bodies). That is the reason why the MRA includes the key comparison made by RMOs in which many non-signatory states are involved. In order to extend the MRA even to non-member states as well as so-called regional economies, the 21st CGPM made a resolution to introduce a new scheme of "Associate" with less responsibility (the minimum contribution is 0.05%) than regular members. The institutes in Associates may be registered in the MRA through RMO key comparisons and the delegates of Associate may attend the General Conference as observers. Now there are 15 Associates, 10 of which already signed the MRA.

(4) Fig. 8 illustrates the position of legal metrology in the whole system of metrological assurance. Because of the trend towards the deregulation, the weight of voluntary calibration is becoming greater and greater in many advanced states and economies. In line with it, it may be assumed that the legal system will be more and more relying on the voluntary system in the future. Of course, the authority responsible for legal metrology is either a part of national metrology institute or a different organization according to the history and structure of each government. But, regardless of the governmental structure, a closer cooperation between the sector of legal metrology and the NMI will be required in the near future, because the present situation of legal metrology may change according to the progress of deregulation as is explained above. The author believes that the CIPM and the OIML should collaborate more closely and more efficiently. To conclude, it may be expected that legal metrology which is responsible for safety and fairness of the public has to and will make use of the scheme of the CIPM/MRA more closely and more effectively.

Fig. 8 Assurance of Traceability of Measurement



Practical Ways in Establishing Traceability in Chemical and Other Measurements

Yoshito Mitani

Centro Nacional de Metrologia, Mexico
Chairman of the Professional Development Committee, SIM
Apdo. Postal 1-100, Col. Centro, Queretaro, QRO. 76000, Mexico
E-mail: ymitani@cenam.mx Phone: (52) 442-2110560; Fax: (52) 442-2110569

Abstract

Practical ways to establish traceability in chemical measurements are examined to understand such diversified measurements, as chemical analysis and testing for conformity assessments. The importance of establishing a conformity assessment infrastructure based on a sound traceable national metrology system is discussed, which should based on the comparable measurement and calibration capability of an NMI.

The importance of chemical measurements and certified reference materials (CRM) for establishing traceability in legal metrology, is discussed taking into account the needs detected in the regional metrology organization in America, the Inter-american Metrology System (SIM). Also, the need of establishing sectorial reference laboratories is discussed.

1. Introduction

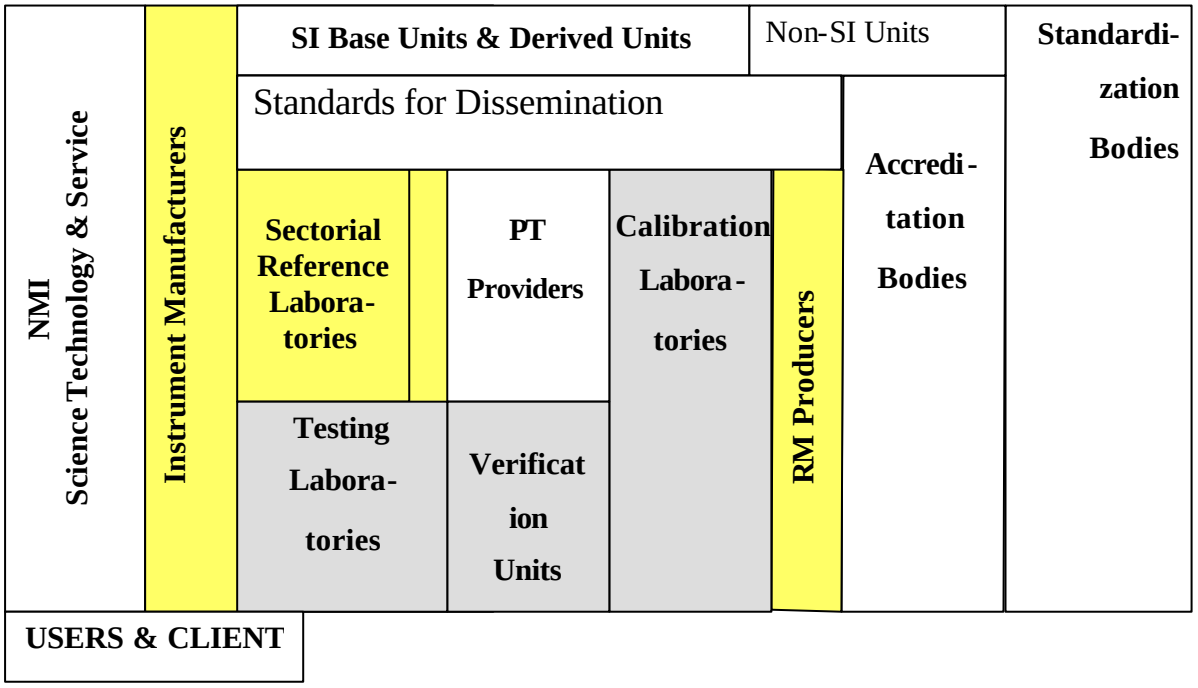
Most of the primary measurement standards are the realization of the SI units, usually under the custody of each country´s National Metrology Institute (NMI). The link between the realization of the SI units and primary standards is established either through primary methods of measurement or calibration of the measurement standards. These are methods which do not require any reference of the same quantity. Additionally, through a series of comparisons between NMIs, comparability of measurements among traceable measurement systems at international level are recognized by each country.

Following the worldwide effort to harmonize measurement capabilities among countries, as a consequence of the strong tendency of globalization of economies, the importance of implementing the traceable chemical and other measurements has been recognized as one of the principal tasks of any NMI.

After the signing of the CIPM MRA in 1999, many NMIs including CENAM has devoted a lot of effort to demonstrate its calibration and measurement capability (CMC), which is required to establish a comparable and internationally recognized national measurement system. In Appendix B of the CIPM MRA the results of Key Comparisons organized by the Consultative Committees

of CIPM are compiled and made available at the BIPM Website http://www.bipm.org, which are considered as supporting evidence of metrological services listed in its Appendix C declared by each NMI and approved by regional and inter-regional review. Another requirement is to support routine provision of those services by a quality system in operation in each NMI.

All elements required to establish a sound metrological system in a country are summarized in the scheme below, and should be functioning in a systematic way so that the traceability chain from the SI units to the final users can be established coherently. In this article, an attempt is made to emphasize the needs and to describe the functions of some of these elements from the point of view of metrological control, and the needs of interaction of NMI with these organizations



Those activities related with metrological assurance and legal metrological control could be organized little bit different among these entities and organizations, if the legal units of measurements are referred to the non-SI units, where the role of the sectorial reference laboratories and the reference materials (RM) producers will be rather critical along with NMI, in order to define a similar scheme of traceability, when it is referred to the SI units.

2. Traceability in chemical and other measurements

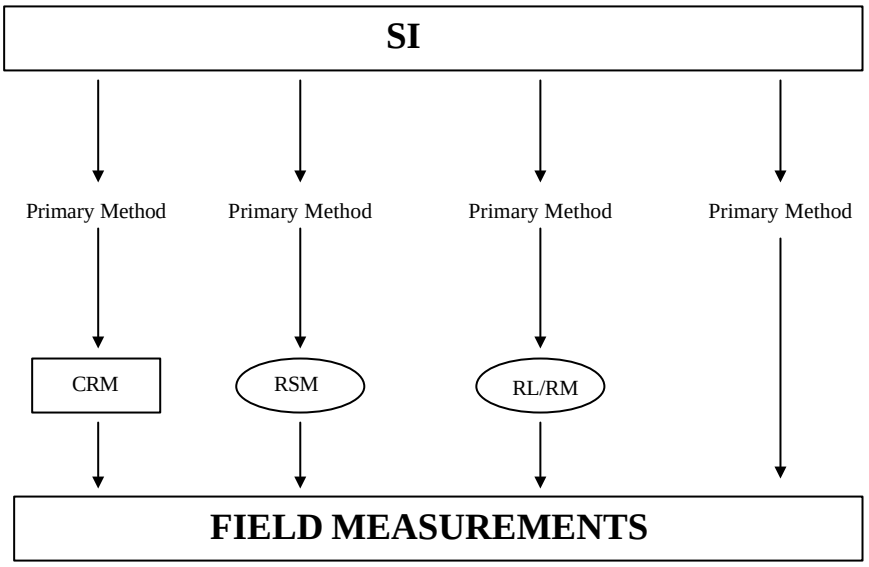
The definition of traceability according to the International Vocabulary of Basic and General Terms in Metrology, is given as follow:

”Property of the result of a measurement or the value of standard whereby it can be related to stated references, usually national or international standards, through an unbroken chain of comparisons all having stated uncertainties”

The traceability definition also can be interpreted according to [1] as follows: "a traceability chain is a chain of values linked by measurements which consist of comparisons of one value, ending in the comparison with the value of the unit we have chosen to express the result of our measurements", with of course all comparisons having stated uncertainties. This interpretation gives clarity in the meaning of traceability concept.

The main parts that support the traceability in chemical measurements are: primary analytical methods, reference materials or instruments and analytical methods validated for specific analyses with specific measurement principle, according to their nature, range of measurements in a specific matrix. These elements should serve to establish an uninterrupted chain of comparisons in chemical measurements and its uncertainty estimation.

As a natural process, and in order to apply the traceability concept for chemical measurements, for which two illustrative proposals have been recognized; to establish a traceability structure which can be set up locally, regionally or internationally, by describing the organizational scheme in a clear and general way as well as its application [1], and the other, to illustrate practical ways of establishing traceability of chemical measurement to SI units by indicating intermediate reference points and primary methods [2]. It is now well understood and widely accepted that a general scheme of traceability must enable one to represent the connection between the definition of the SI units and the results of a measurement in terms of SI units, by a series of measurements at intermediate reference points, which may be realized by reference materials, reference instruments or reference measurement methods maintained by reference laboratories. The definition of the SI units are realized by primary methods at the highest metrological level. This scheme is shown in figure 1.



CRM:Certify reference material
RSM: Reference system of measurement
RM:Reference method
RL: Sector Laboratory of Reference

Figure 1

Then, the dissemination of the accuracy of the standards can be established in all the chemical and other measurements by the application of one of the following mechanisms [2]:

i. Use of reference materials traceable to SI. In the majority of the measurements, the certified reference materials (CRM) traceable to SI are by far the best definable reference points and they are most frequently used as measurement standards in chemical and other measurements. These materials are the means of achieving reliable measurements and they are available from the internationally recognized organizations, although the distribution is not necessarily enough for a wide range of users.

ii Reference Systems of Measurement. This route of traceability is based on the measurements made with a measurement system or instruments developed to create reliable intermediate reference points for measurements; these systems are calibrated against measurement standards traceable to SI units, by means of a thermodynamically acceptable reference state.

iii Reference measurement methods, which produce values traceable to the SI units, when the recommended rules of procedures are applied. Most measurements belong to this group, because there is still limited the availability of CRM and reference system for chemical and materials properties measurements. These measurements shall be carried out by laboratories that have demonstrated competence in applying these measurement methods supported by the SI units. There are, however, many method dependent measurements described by non-SI units in the field of legal metrology, for which more attention should be paid in order to make these measurements comparable through the establishment of a objective and scientifically demonstrable traceability chain.

iv Primary methods applied directly to routine measurements. This route corresponds to the cases in which a analytical laboratory is able to establish primary methods to establish a direct link between their measurements and the SI .

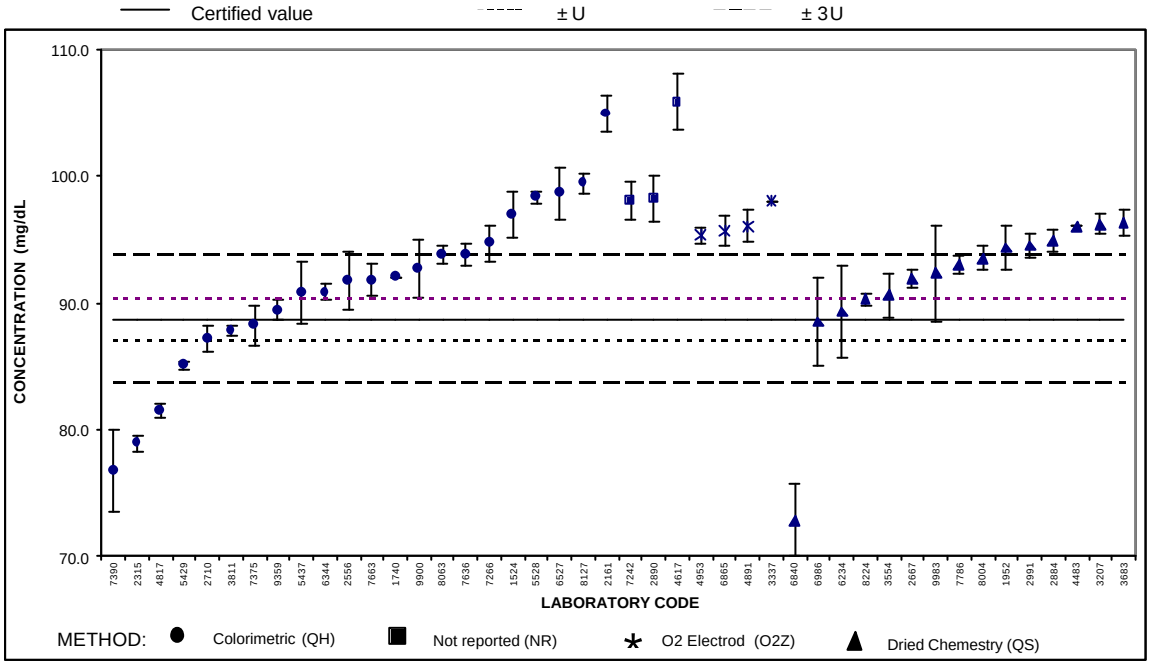
3. Traceability and Legal Metrology

In the context of legal metrology, the above mentioned mechanisms have been supported by corresponding actions, such as

i) Declaration of certified reference materials for state use
ii) Type approval of measurement devices
iii) Reference methods for mandatory application
iv) National recognition as primary standards
v) Authorization of traceability to higher order measurement standards of other countries, if there is no such higher order reference in a country

In new fields of metrological interest, such as In Vitro Diagnostic (IVD) medical devices [3], where measurement of quantities in samples of biological origin is involved, the metrological traceability of the value assigned to calibrators and control materials is identified by the traceability chain and calibration hierarchy.

Fig. 1a Mean values and standard deviation for Glucose concentration in sample 1



CERTIFIED VALUE: 88.7 mg/dL
Uncertainty (U): ± 1.68 mg/dL

The correct application of the traceable measurements is essential for the establishment of confidence on the decision making for legal actions. On the other hand it is also essential for the establishment of confidence on conformity evaluation which has economic impact.

These are the critical part of this kind of products: to ensure the clinical diagnosis based on the traceable value and to conduct sound type approval to eliminate non conformity products, based on the traceable measurements.

A good example is illustrated in figure 1a, which shows results of a proficiency testing on glucose measurements in human serum, in which some 44 clinical laboratories participated with commercially available IVD devices. The organizer presented the results and compared with the certified value, denoted by the solid line in the figure, of the reference material (CRM) employed as the test material. In this figure the results are grouped in 4, according to the methods applied, or equivalent to say, according to the commercially available kits. The organizer considered that the group identified as QH has deficiency in homogeneity, while the groups O2Z and QS show significant bias with respect to the certified value and to the uncertainty level offered by the organizer **[5].**
Assuming that this result might raise a non conformity to the acceptance level of IVD kits against these two manufacturers, who claim the traceability to the respective national standard, the issue turns out to be the degree of equivalence of the national standards among countries, because the manufacturer claims traceability to his respective national standards. In this particular case, if this methodology is applied to the conformity assessment of a measuring instrument or device, the manufacturer could claim for the validity of the value of the CRM used for this study or the for the demonstrated capability of the institution responsible for assigning such value.

This is one of the strong arguments that international comparability should be achieved based on the comparisons among national standards of respective country, which are carried out periodically by Consultative Committees of BIPM, who maintains Key Comparison Data Base as supportive evidence of the degree of equivalence between countries.

4. Elements of traceability in Legal Metrology

According to the recent survey done by Legal Metrology Working Group of SIM for training needs, we can recognize some commonly required quantities of measurements. To support the legal metrological control, it is meant principally the legal control of measuring instruments, such as type evaluation and approval, verification and inspection, as well as the metrological supervision and metrological expertise. However, it is now well recognized that the measurements required in legal metrology is not only for ensuring the equity in commerce and trade, but also for establishing reliability for decision making in health care and environmental protection, where more emphasis on the measurements of amount of substance should be made and consequently on certified reference materials (CRMs) and methods of measurements, as it was revealed in the recent survey among SIM countries on the needs for development for human resources.

The most common fields of interest in development of human resources legal metrology in SIM counties are referred to measuring instruments, such as:

- large capacity meters in fuel storage terminals,
- small capacity scales (<25kg) in supermarkets, grocery stores, and other markets,
- gasoline and diesel fuel dispensers in service stations,
- fuel oil and gasoline meters (medium-capacity meters) on trucks, medium and large capacity scales (500kg to 100 000kg);
- water and gas meters,
- automatic rail-weighbridges and automatic instruments for weighing road vehicles in motion,
- sacales for measurements and verification of spray net content, cylindrical liquid gas of net content
- electric energy meters;
- sphygmomanometer and clinical thermometer,
- sound level meters.
- taximeters

Nevertheless, more diversified requirements are found in chemical measuremnts where there is the need to establish reference materials, such as;

- CRMs for pH, electrolytic conductivity for water quality.
- CRMs for pollutants in air, water and soil
- CRMs for gass components in vehicle exhaust emissions

This shows that legal metrological control covers not only measuring instruments, but also some specific CRMs. In this context we suggested the use of the word CRMs in the fields covered by the state metrological control in OIML TC 3/SC 3 in 1999, which was approved. In this document the concept of type approval of CRM is introduced, which gives the equivalent sense of supporting the metrological control, sincer CRM is used as a reference point of measurements, whose issue is not an one-time occurence, but shall be renewed. These two elements are considered equivalent in establishing traceability; namely they correspond to the cases i) and ii) of the section 2, respectively.

5. Use of CRMs in testing laboratories

Most of the testing laboratories need to demonstrate the validity of measurement methods, particularly in analytical laboratories where separation techniques are the main difficulties, because it is not easy to establish the traceability chain by the use of calibration standards of simple matrix. Consequently, either the validation of analytical methods or calibration by complex matrix reference materials is required. However, unless the process is clearly described with corresponding uncertainty, the validation process becomes a bottleneck for establishing a traceable measurement. Then, in most applications, the role of CRMs of a similar matrix becomes crucial in the quality of measurements.

CRMs are also important for testing laboratories other than those performing chemical measurements. Hardness and toughness are some of the best known mechanical tests which

require CRMs. The role of these CRMs are precisely to validate the testing methods and to calibrate the instrumental set-up.
In the strict sense of measurement traceability, the use of CRMs in validation processes do not constitute a part of traceability chain, but they are essential for the estimation of the uncertainty of the methods used, which give the traceable measurement.

However, there are some other important issues to be noted. Based on our recent assessment of traceability of commercially available RMs, we are recommending to testing laboratories and commercial suppliers of chemicals to distinguish clearly CRM quality products with suitable certificates from other chemicals and reagents dedicated to other uses. This assessment has been requested by authorities and is now under practice as a part of the formal recognition process of accredited testing laboratories that have to demonstrate their capability to conduct traceable measurements through usage of standards traceable to national standards of foreign countries, instead of national standards, in case they are not available in the country.
During the course of assessment, it was found that commercial suppliers of chemicals normally do no declare uncertainty, and if they declare it, it is too small to be regarded as reliable without any additional supporting evidences.

On the other hand, in order to combine the capability of these commercial producers and the capability of the certification of NMI, CENAM has launched a program called Certified Traceable Reference Materials, MRTC for its acronym in Spanish, similar to the NTRM initiative of NIST in USA.
This program is intended primarily to promote the capability of domestic industries to produce and certify CRMs in those fields where there exist enormous demand and absolute lack of them.

The scheme of a national metrology system presented in the section 1 illustrates these important elements like instrument manufacturers and RM producers. The role of NMI in coordination with the organization of state metrological control is to promote these elements according to the needs of the society. In any case the important condition for the success of the mechanism is the scientific and technological competence of those organizations, which support the economic development of a country. In Fig. 2 it is schematically shown how the interrelation between nation standardization bodies and metrology infrastructure should be organized to support harmoniously national conformity assessment infrastructure.

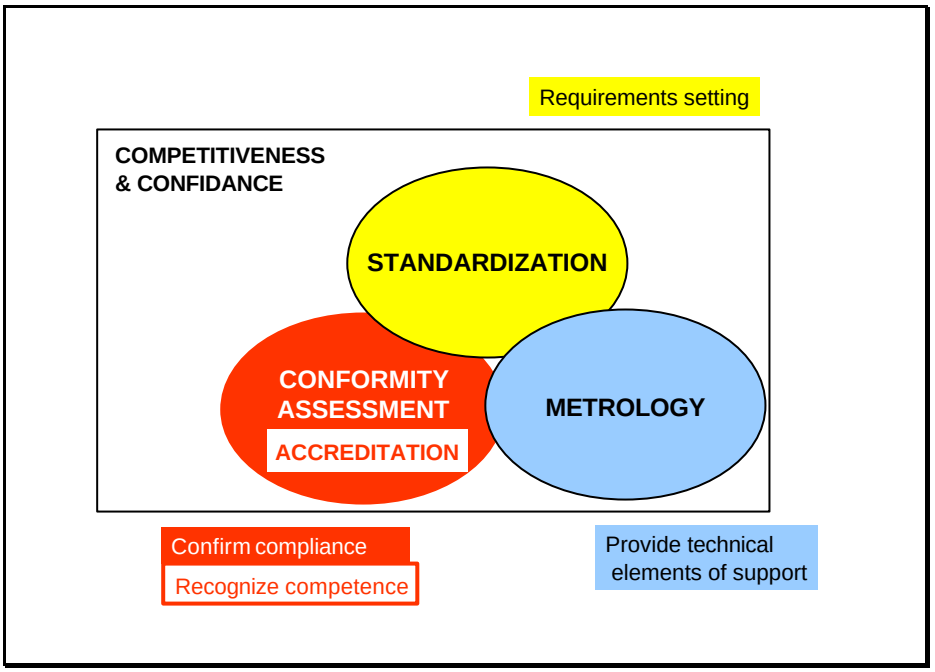


Fig. 2 Elements forming a national conformity assessment infrastructure.

This is the rationale to look for a coherent and harmonious regulation practice, and consequently we need to establish a consistency in the legal and scientific metrology. In fig. 3, a schematic organizational chart to show how an NMI can provide traceability to the whole national metrology infrastructure, and how they should interact each other, considering all elements cited in the schematic figure in the section 1.

The role of reference laboratories for the establishment of traceability is especially important in legal metrology, because it is very often specified the use of those legal units which do not belong to the SI units. So far the traceability is meant to the SI units, which have universal acceptance thanks to their rigorous scientific basis, while other non SI units require special considerations and are usually method dependent. The development of methods for non SI units depends also on the development of scientific knowledge on the measurand and technological capability to identify and quantify it, based on the well defined measurement principle and its technological realization. The achievement of these efforts will turn out to be either the development of new methods of measurements, new instruments or devices and possibly CRM´s, which are basically three reference points of those mentioned in the section 3 for chemical and other measurements.

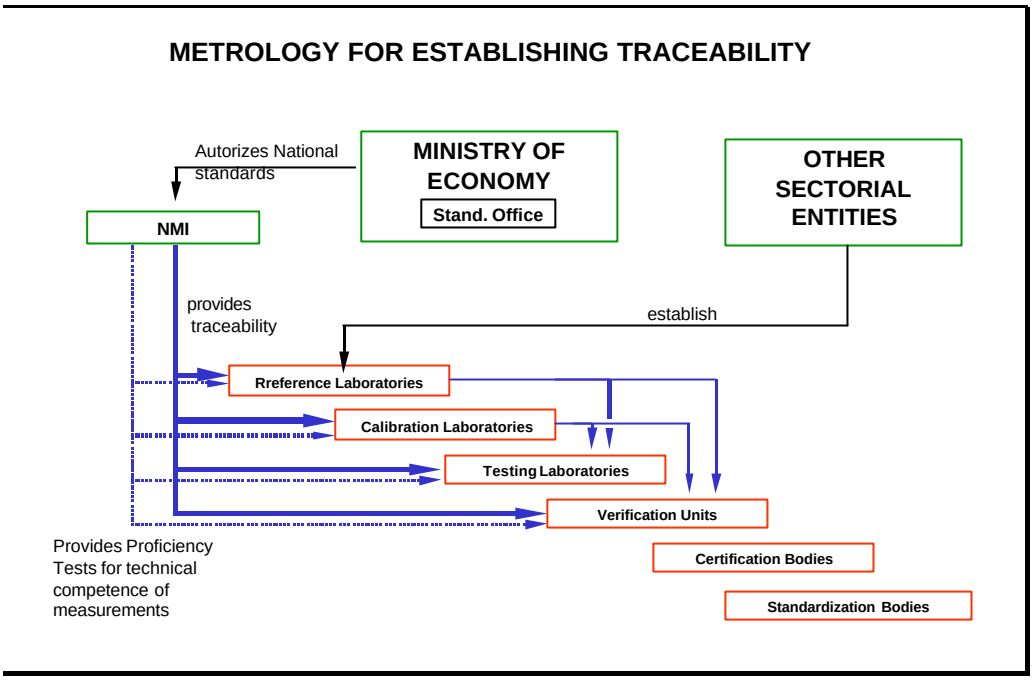


Fig. 3 Establishment of traceability of a conformity assessment infrastructure

6. Promotion of reference laboratories?

When there is no CRM or other measurement standard to provide traceability, this should be established through metrologically sustained methods.
As was mentioned previo usly, most measurements in chemical analysis and testings are method dependent and it may be necessary to recognize the laboratories that have competence and capability to carry out measurements with demonstrable traceability to SI. Consequently reference laboratories maintain reference procedures, which provide a reference values to routine measurements. The competence of reference laboratories with regard to environmental conditions, staff, equipment, use of suitable procedures and management can be subjected to an assessment within the framework of an accreditation process. A reference laboratory shall demonstrate its competence to apply the reference methods properly. The reference method shall be validated. and also verified on the basis of documented reference procedures and on the results of parallel comparative measurement [4], in which the participation of NMI is essential to emphasize metrological robustness of the method.

In this context, complex reference materials, where they are technically feasible, are normally developed by NMIs for the validation of reference procedures, and help laboratories establish traceability of their measurements in the sense that these measurements are supported metrologically by traceable measurements to the SI units, and are capable of reproducing the value within the acceptable uncertainty. The so-called reference laboratories are expected to be capable of conducting the validation process along with NMI.

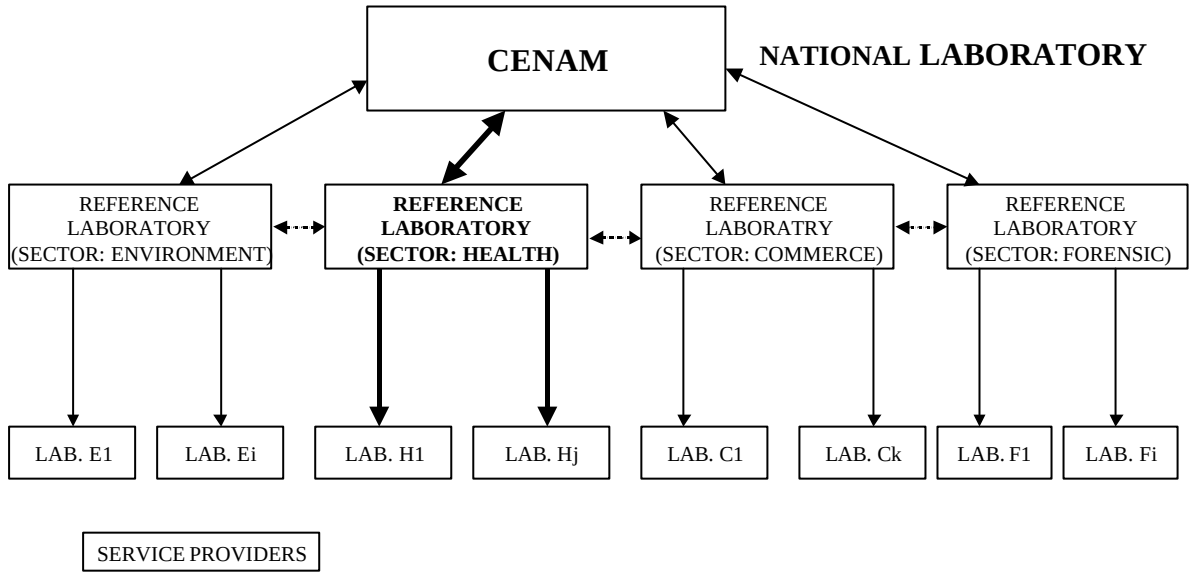
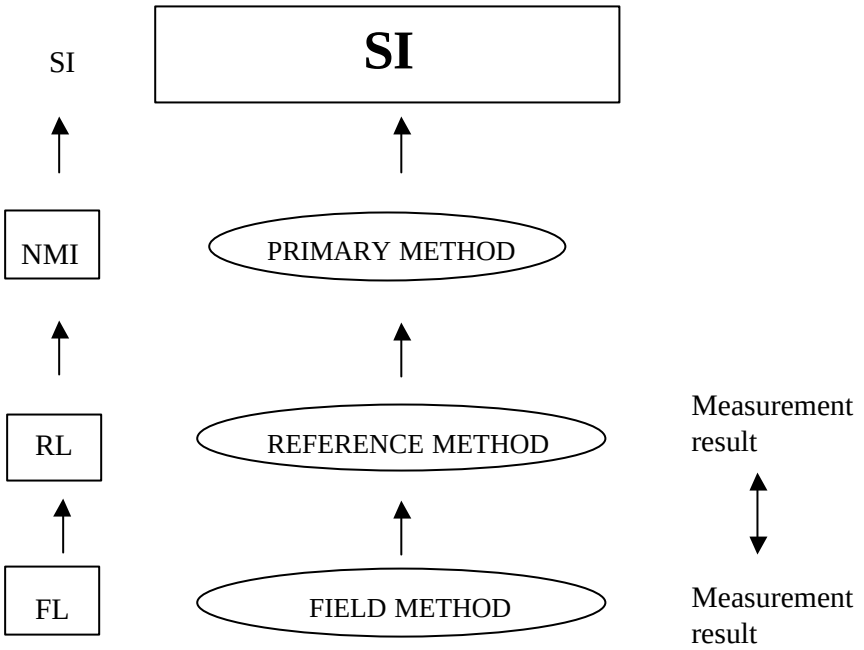


Figure 4

Method-dependent measurements can be grouped by sector. For example, in the clinical fields there are cases where some higher order reference materials are required for IVD methods, such as the determination of glucose in human serum. Reference laboratories are also required for specific measurement methods. These issues are now under the responsibility of JCTLM (Joint committee on the Traceability of Laboratory Medicine of CCQM). One of the important tasks for NMI´s in collaboration with legal metrology service entity is to develop reference laboratories in a country in each sector of importance. The metrological scheme for the definition of reference laboratories can be represented in figure 4.



In order to have reliable and acceptable measurement results, the traceability to SI quantities (kg, s, A, K, mol, cd), involved in the procedure of measurement of the reference method and field method, must be assessed by mean of a calibrator and a series of comparisons between NMI and RL and, RL and FL.

Figure 5

7. Sectorial Reference Laboratories for legal metrology

It is considered necessary to involve all governmental entities that have the responsibility in conformity assessment to regulations in the establishment of sectorial reference laboratories. From this standpoint, more collaboration is expected between NMI and public sectorial laboratories, which are the technical authorities in the surveillance of mandatory standards and regulations. The idea is to give them a metrological responsibility in that sector called sectorial reference laboratories. Their functions are expected to be as follows:

- Establish traceability of their measurements to NMI in all the quantities required in their field of responsibility
- Provide Proficiency Testing (PT) to field laboratories to establish comparability and reliability
- Develop and validate analytical methods in the field of responsibility
- Conduct type approval of measurement instruments used by the field laboratories for the conformity evaluation to the specific regulations under their responsibility

It is intended to share metrological activities between NMI and reference laboratories responsibility in their respective level and fields, by maintaining coherent and comparable measurement capability among reference laboratories and consequently providing traceability to field laboratories, see Fig. 5.
These tasks may deserve the highest priority of the government in the next few years, to extend collaboration programs to the fields of pharmaceuticals, clinical, health, environmental, agricultural and forensics, and also to look for modifying some part of the actual law to assign explicitly the specific metrological responsibilities to these reference laboratories.

8. Final remarks

From the practical point of view, a conformity assessment infrastructure is considered which is organized in such a way that it could promote the dissemination of the accuracy of the units realized in national standards. In order to confront the extremely diversified and demanding tasks required, it is suggested to identify a series of sectorial reference laboratories which could take part in the metrological responsibility with NMI, particularly for legal metrological control.
The availability of measurement instruments and CRMs depend on the degree of economic development and NMIs should take lead the promotion of those organizations who can participate in such activities as to develop CRMs under a scheme which establish traceability to the SI, while developing comparable capability in measurements and calibration according to the CIPM MRA requirements.

9. Acknowledgment

Thanks are due to Mr. Felix Pezet, CENAM for his valuable comments on Legal Metrology, and also to Mr. Cesar Luiz da Silva, INMETRO, chairman of the Legal Metrology Working Group of SIM for the information on the training needs. Fruitful discussions with Mr. Alejandro Perez, Ruben Lazos and Dr. Melina Perez of CENAM are highly appreciated.

References

[1] De Bièvre P. The key elements of traceability in chemical measurement: agreed or still under debate?. *Accred. Qual. Assur.* 2000; 5:423-428

[2] Richter W. Primary methods of measurement in chemical analysis. *Accred. Qual. Assur* 1997; **2**:354-359

[3] In Vitro Diagnostic Medical Devices-Measurement of Quantities in Samples of Biological Origen-Metrological Traceability of Values Assigned to Calibrators and Control Materials, pr EN ISO 17511 (Geneva : International Organization for Standardization, 1999).

[4] Siekmann L. Measurement traceability in clinical chemical analyses. *IVD Technology Magazine* 2002

[5] M. Perez-Urquiza, et. al. Use of reference value for the Proficiency Test scheme in the measurement of glucose and cholesterol in human serum., Accred. Qual. Assur. 2003 to be published

Legal Traceability of Breath Alcohol Measurements

Judith Bennett, Dr

National Standards Commission

PO Box 282 North Ryde, New South Wales 1670, Australia

On behalf of the project team:

Dr Laurie Besley (National Analytical Reference Laboratory)

Dr Richard Brittain (National Standards Commission)

Mr John Farrar (Victoria Police Laboratory)

Ms Yen Heng (National Standards Commission)

Professor Brynn Hibbert (University of NSW)

Dr Lindsey Mackay (National Analytical Reference Laboratory)

Dr Brett Yeomans (Australian Government Analytical Laboratories)

The legal metrology system in Australia

In Australia, measurements which are made for any legal purpose must comply with the provisions of the National Measurement Act 1960 and the National Measurement Regulations 1999. There are three key elements that need to be addressed in order to ensure that measurements can be defended in court under the provisions of the National Measurement Act:

i. Measurements must be expressed in the Australian legal units of measurement prescribed for the particular physical quantity being measured.

ii. Measuring instruments must be of an approved pattern, as evidenced by a pattern approval certificate issued by the National Standards Commission.

iii. Measured values must be legally traceable to an appropriate Australian standard of measurement, as evidenced by a certificate of verification or certification, issued by an appropriate authority, for each link in the chain of measurements.

An important (and possibly unique) aspect of the Australian legislation, is the certificate system. Certificates issued under specific Regulations by appointed authorities are generally accepted as evidentiary instruments in a court of law, thus avoiding the 'expert witness' approach to litigation.

The application of legal metrology to breath alcohol measurements

To date, the focus of legal metrology in Australia has been on trade measuring instruments. However, there is increasing recognition of the need for regulators and law enforcement authorities to have a sound evidentiary basis for measurements on which legal prosecutions may depend. For some time, police services have been vulnerable to legal challenges of their breath alcohol measurements, and this situation prompted them to seek help from the National Standards Commission.

In a collaborative project between the NSC and all police services in Australia, systems have been put in place to enable police officers to make breath alcohol measurements which comply with the requirements of the National Measurement Act.

Units of measurement

Driving offences are specified in State and Territory legislation. Offences for driving under the influence of alcohol have historically been based on blood alcohol content, defined in the legislation as: grams of alcohol in 100 millilitres of blood. Analytical laboratory techniques have been developed, which make it possible to measure blood alcohol content with considerable accuracy. Whilst blood tests are useful for forensic purposes, or as the basis for an appeal, they are not practical for routine roadside testing of drivers. In particular, the taking of blood samples requires the consent and cooperation of the driver, and the results are not instantly available.

Breath analysis (patented as the Breathalyzer) was developed in the USA in the 1950s and was used in most States in Australia by the early 1970s. Initially, breath analysis was a chemical process, based on titration with potassium dichromate. This cumbersome process was replaced in the early 1990s with electronic breath analysis, based on the absorption of infra-red light (at selected wavelengths) by alcohol in a sample of air.

Although breath analysers actually measured the amount of alcohol in a sample of air, the result was nevertheless recorded as grams of alcohol in 100 millilitres of blood. From a legal metrology perspective, this is a clear example of a measurement which is not what it purports to be. The problem lies in the use of a “conversion factor”.

When evidential breath analysers were introduced, Australian police followed the practice of police in the USA, and used a factor of 2100 to convert breath alcohol content to blood alcohol content. This conversion factor was based on research conducted from 1950 onwards, and was adopted by Australia in 1961. However, the partition of alcohol between blood and breath in the human body is not constant across the population. *In vivo* determinations of this

conversion factor are derived from statistical analyses and are susceptible to many influence factors.

This indirect method of measuring blood alcohol content is clearly not acceptable legal metrology practice. The decision was therefore made to amend the legislation in all jurisdictions, to include an offence based on breath alcohol content, defined as: grams of ethanol in 210 litres of exhaled breath. This definition was adopted in order to retain the same nominal value for the legal limits, which had strong public recognition. It meets the requirements of the National Measurement Act, as base SI units are Australian legal units of measurement.

Pattern approval

The National Measurement Act gives the National Standards Commission the authority to examine and approve the patterns of measuring instruments used for trade and for any other legal purpose. Pattern approval requirements are based on OIML Recommendations (where they exist), but may be modified in consultation with stakeholders, to account for climatic conditions in Australia, and any other local requirements which do not breach the WTO Technical Barriers to Trade Agreement.

OIML R126 for evidential breath analysers, was published in 1998 and formally sanctioned by the CIML in 2000. A working party with representatives from all Australian police services used this document to develop the Australian pattern approval specifications. The OIML document provides alternative test methods using either dry or wet gas. The Australian standard NSC R 126 is based on the wet gas method. It was accepted by all stakeholders in 2000, but its implementation has taken some time, owing to the need to design, manufacture and commission the complex lung simulation equipment required for pattern approval testing and initial verification of instruments.

The lung simulator and associated equipment were jointly funded by the police jurisdictions and reside within the Technical Services Laboratory of the Victoria Police for the purposes of some of the pattern approval testing. The first pattern approval applications have been received. It is anticipated that pattern approval will become mandatory in early 2004.

Legal traceability to an Australian standard of measurement

The final, and most challenging task has been to ensure that police officers can make breath alcohol measurements (using pattern approved evidential breath analysers), such that the measured values are legally traceable to Australian primary standards of measurement.

As evidential breath analysers are verified and calibrated using aqueous ethanol solutions, the challenge was to supply aqueous ethanol solutions that would meet the Australian legal definition of a "certified reference material". This means a reference material certified under Regulation 48 of the National Measurement Regulations 1999, by a Certifying Authority appointed by the National Standards Commission.

The provisions of the regulations and conditions of appointment of certifiers are such that they encompass the VIM definition of a certified reference material:

> *"a reference material accompanied by a certificate, one or more of whose property values are certified by a procedure which establishes traceability to an accurate realisation of the unit in which the property values are expressed, and for which each certified value is accompanied by an uncertainty at a stated level of confidence"*

but additionally ensure that property values are traceable to Australian legal units of measurement (SI units); that certifiers hold appropriate national standards of measurement and are accredited to ISO Guide 34; and that certificates are issued which are accepted as evidentiary instruments in Australian courts of law.

Collectively, Australian police services use approximately 1000 litres of aqueous ethanol solution per annum for calibration of breath analysers. In the past, these solutions were produced by gravimetric dilution of analytical grade absolute alcohol with water, against calibrated masses. Considerable care in drying and handling ethanol is required to produce consistent results. However, simple gravimetric preparation of solutions can never confer properties of traceability owing to the hygroscopic nature of ethanol. Some producers of aqueous ethanol solutions determined the mass concentration of their solutions by oxidation with excess potassium dichromate but this method, in turn, relied on knowledge of the purity of the potassium dichromate. In the absence of a potassium dichromate solution traceable to SI units, this method did not satisfy the legal requirements.

An alternative approach was needed, and assistance was sought from colleagues at the National Analytical Reference Laboratory (NARL). Coincidentally, NARL had recently developed a primary method for the determination of ethanol mass fraction in aqueous solutions, using isotope dilution mass spectrometry (IDMS). These measurements were directly traceable to the Australian kilogram. NARL had also participated in an intercomparison of measurement standards for ethanol in aqueous matrix, conducted by the Consultative Committee for Amount of Substance (CCQM-K27), and their results agreed well with the gravimetric values used as the benchmark for the study. The intercomparison therefore provided an important validation of the methodology.

IDMS is an expensive process and is not an affordable method for the routine measurement of working standards of ethanol. It was therefore decided that IDMS would be used to characterise a 'national standard' CRM which would then be used to calibrate the apparatus used by the manufacturers of the 'working standard' solutions supplied to the police. This approach gives the manufacturers the option of using potassium dichromate titration or gas chromatography to analyse the working standard solutions they produce. All of the intermediate steps in the traceability chain are illustrated in Figure 1 (a) for the titration method, and Figure 1 (b) for the gas chromatography method.

Figure 1 also indicates typical assigned values and uncertainties for each measurement step. The uncertainty assigned by NARL to the CRM is better than 1.0% (at the 95% confidence level, U_{95}). The uncertainty expands with each step in the traceability chain, but ultimately enables police to measure the alcohol content of a motorist's breath with an uncertainty of about 6% (U_{95}).

The national standard solutions are produced by the Technical Services Laboratory of the Victoria Police and are currently certified by NARL for a period of 6 months. A separate research project is being undertaken to investigate the long-term stability of these solutions, as a basis for reviewing the certification period. The CRM solutions are individually sealed with a tamper-evident device and delivered in a locked transit case (for chain of custody purposes) to laboratories that produce 'working standard' solutions for the police. The CRM solutions are accompanied by safety data sheets and recommendations for storage and use. Bottles are individually serial-numbered and intended for single calibration only. Police officers have been supplied separately with detailed instructions on the calibration and operation of the evidential breath analyser.

NARL and the two laboratories that manufacture the working ethanol standards have been appointed as Certifying Authorities, so that they can issue certificates which (collectively) provide evidence of traceability, if needed in court.

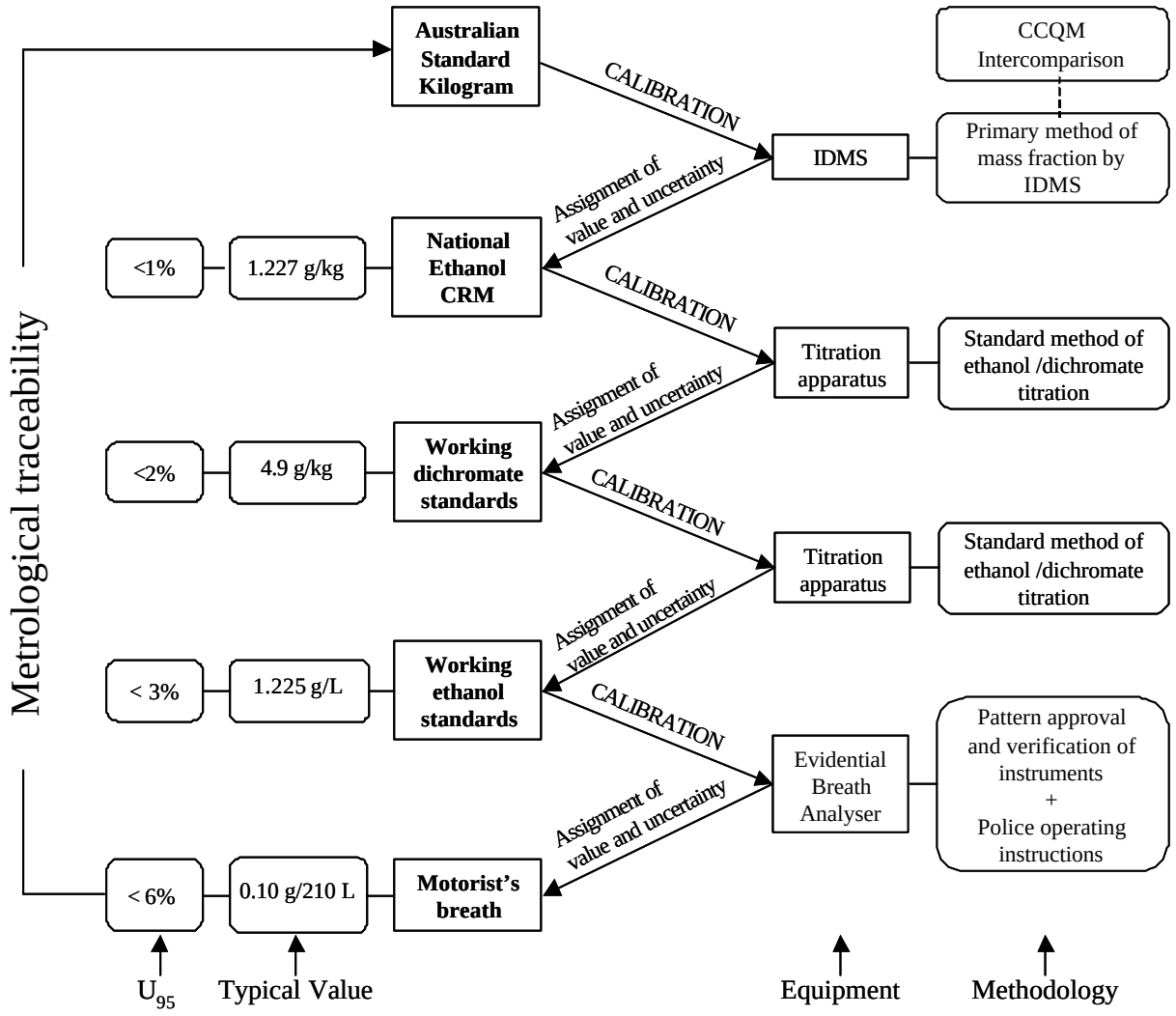


Figure 1: Legal Traceability of Breath Alcohol Measurements

(a) using the titration method to calibrate working ethanol standards

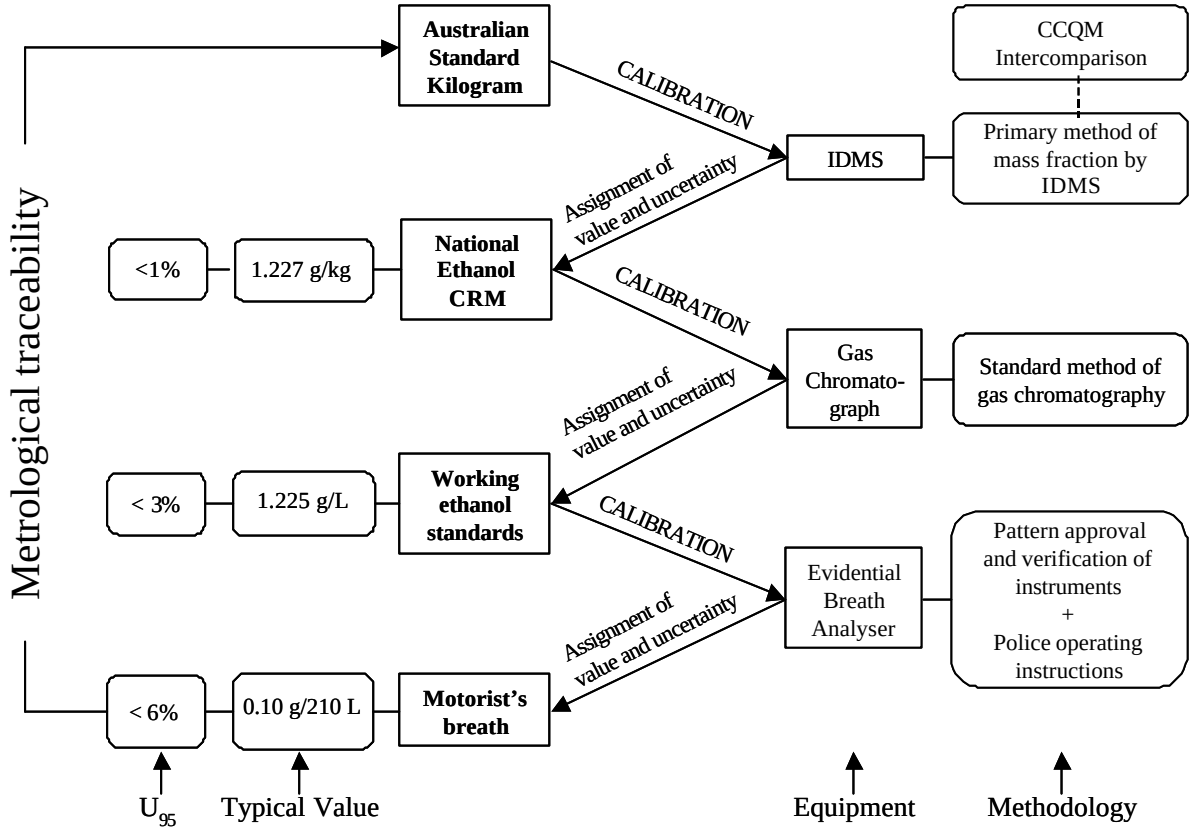


Figure 1: Legal Traceability of Breath Alcohol Measurements

(b) using gas chromatography to calibrate working ethanol standards

Conclusions

It is now possible for police officers in Australia to make breath alcohol measurements which comply with the requirements of the National Measurement Act and are therefore much less likely to be successfully challenged in court.

This goal has been achieved through the close collaboration of physical, chemical and legal metrologists, manufacturers of instruments and chemicals, and the police services who are the clients and ultimate beneficiaries of this work.

This project has demonstrated, for the first time in Australia, that legal metrology concepts can be successfully applied to chemical measurements. At the same time, it has illustrated the complexity involved in establishing traceability in chemistry – even for a relatively simple aqueous ethanol solution. It is an indication of the challenges that must be faced in the future, if legal metrology requirements are to be applied to a broader range of chemical measurements. Environmental regulations, for example, are concerned with the measurement of small traces of contaminants in complex matrices such as soil, river water and air. We have taken the first step, but we have a mountain to climb!

Traceability of Mass Measurements for Legal Metrology in Malaysia

Soo Fatt Chen

National Metrology Laboratory, SIRIM Berhad

P. O. Box 7035, Section 2, 40911 Shah Alam, Malaysia

Abstract

The establishment of a credible system of measurements with international traceability is vital to the economic and social development of a country. This paper describes the traceability system of mass measurements in Malaysia for legal metrology. It traces the development of the weights and measures system in the 1960s to the present which is based on the International System of Units (SI). The hierarchy of mass standards as prescribed under the Weights and Measures Act 1972 and the manner in which the international traceability of the Malaysian kilogram standard is achieved are described.

Introduction

Malaysia has a total land area of approximately 329,847 km^2 and a population of 24.4 million based on a population census undertaken in 2000. It consists of fourteen (14) states administered by a Federal Government. The country is made up of two regions separated by about 540 km of the South China Sea; Peninsular Malaysia (West Malaysia) which is situated at the edge of continental Southeast Asia and the states of Sabah, Sarawak and the Federal Territory of Labuan (collectively referred as East Malaysia) are situated in the north-western coastal part of Borneo Island.

The weights and measures system of Malaysia in the 1960s and earlier was rather disorganized. It consisted of different systems individually administered by the various states under their respective weights and measures enactments. These different systems were based on a multiplicity of units such as the British "imperial" units, Chinese units, metric units, indigenous or customary units and others whose origin was obscure. The values of some units varied from place to place within Malaysia, and the values of denominations within series of weights and measures often incurred a confusing progression which hindered efficient usage.

Figure 1 : Map of Malaysia

In 1972 Malaysia launched a 10-year 'metrication' programme to harmonize the national weights and measures system with the International System of Units (SI). This was done through the enactment of a Federal law, the Weights and Measures Act 1972 which enabled the use of the SI system of units to be enforced and to allow for a period of transition between the old systems and the new system. During the period of transition numerous educational campaigns were conducted by the Ministry of Domestic Trade and Consumer Affairs for the public on the use of the metric system. The Act was fully enforced in October 1982.

Figure 2 : 'Imperial' pound standards issued to inspectors of weights and measures in the 1960s (L to R : 14, 7, 4, 2, 1 lb)

Weights and Measures Act 1972

The Weights and Measures Act 1972, as amended by further Acts in 1981, 1990 and 1992 prescribes the present Weights and Measures system in Malaysia. The Act prescribes the establishment of standards of mass and measure based on the SI and the administrative arrangement for implementing the provisions of the Act.

The Act is enforced by the Ministry of Domestic Trade and Consumer Affairs. The officer overseeing the overall implementation of the Act is the Chief Inspector of Weights and Measures who is also the Director-General of the Enforcement Division of the same Ministry.

The original scope of the Act was confined to trade measurements; regulating fair trade practices and control of instruments used for such purposes. It was amended in 1990 to cover other fields of measurements as well. Its existing provisions are however still very much focussed on enforcing legal metrology for trade.

Appointment of a Custodian of Weights and Measures

The Weights and Measures Act 1972 provides for the appointment of a Custodian of Weights and Measures (henceforth referred to as the Custodian) who shall be responsible for the realization, establishment, maintenance and custody of the Malaysian national standards of measurement based on the SI. The National Metrology Laboratory, SIRIM Berhad (NML-SIRIM) has been appointed as the Custodian.

Hierarchy of Mass Standards for Legal Metrology

Mass standards prescribed for the enforcement of the provisions of the Weights and Measures Act are classified as follows; primary standards, secondary standards, tertiary standards and working standards.

Designated organizations to maintain such standards and the means and manner by which the standards are to be verified and certified are prescribed by the Minister responsible for the Act.

Working Standards

Working Standards of mass are kept in the custody of the Inspectors of Weights and Measures and are mainly used for the inspection, verification and certification of weights and weighing instruments used in the market place for direct retail trade. The most common weighing instruments verified by the inspectors are the spring scales and platform weighing scales of the floor and counter types.

The Ministry of Domestic Trade and Consumer Affairs has a total of 50 weights and measures offices located in the various states throughout Malaysia. Each office is provided with a minimum of the following standard weights as Working Standards :

(i) 2 sets of chromium-plated brass or brass weights, 1mg – 20 kg, of OIML Classes F2 and M1 and,
(ii) 120 pieces of 20 kg cast iron weights of OIML Class M2.

Figure 3 : Inspectors' Working Standards, 10 g – 20 kg

Six (6) major branch offices are each additionally provided with 40 pieces of 250 kg roller-type cast iron weights of OIML Class M2. These weights are used for the verification of weighbridges used for trade purposes as well as those installed along the highways for road safety enforcement under the Road Transport Act 1987.

The Working Standards are verified and certified by comparison with tertiary standards once every year by the Custodian. During the verification the weights are adjusted to within the limits of error prescribed for their respective accuracy class.

Tertiary Standards

The Weights and Measures Act 1972 has designated the Custodian (NML-SIRIM) to be responsible for maintaining Tertiary Standards of mass. Copies of these standards are also designated under the same Act to be maintained by two other organizations :

(i) Weights and Measures Offices and,
(ii) Chemistry Department of Malaysia.

The Tertiary Standards maintained conforms to OIML Class F1 weights. The nominal values of such weight range from 1 mg to 20 kg. Thirty six (36) branch offices of weights and measures are each equipped with a set of the Tertiary Standards. These branch laboratories are provided with general air-conditioning.

The Custodian uses the Tertiary Standards in its custody for the verification and certification of the Working Standards as described above.

Tertiary standards maintained by the inspectors are generally used for the verification of precision weighing balances used for the transaction of precious metals. The branch offices are however not adequately equipped and trained to do the verification of the working standards themselves.

The Chemistry Department provides forensic testing services to the police department and other government agencies. The department uses a number of analytical balances in the course of their work. The Department has 10 laboratories located throughout Malaysia. The Tertiary Standards are mainly used for the verification of these balances. Higher precision balances are submitted to NML-SIRIM for calibration.

Tertiary Standards of mass are verified and certified by the Custodian (NML-SIRIM) at intervals not exceeding 3 years.

Secondary Standards

The Secondary Standards are kept in the custody of the Custodian. These standards are stainless steel weights conforming to OIML Class E2. NML-SIRIM at present maintains two sets of such weights from 1 mg to 20 kg. The standards are verified and certified against the Malaysian Primary Reference Masses at the NML-SIRIM at intervals not exceeding 5 years.

Primary Standards

The Malaysian Primary Kilogram maintained by NML-SIRIM is the national standard for the unit of mass in Malaysia. It is cylindrical in shape and made of austenitic stainless steel of density 8000 kg/m^3. The standard conforms to OIML Class E1 and has been calibrated at an overseas laboratory traceable to the International Bureau of Weights and Measures (BIPM) to a relative expanded uncertainty of $\pm$ 1 x 10 $^{-7}$ (k=2, 95% c.l.)

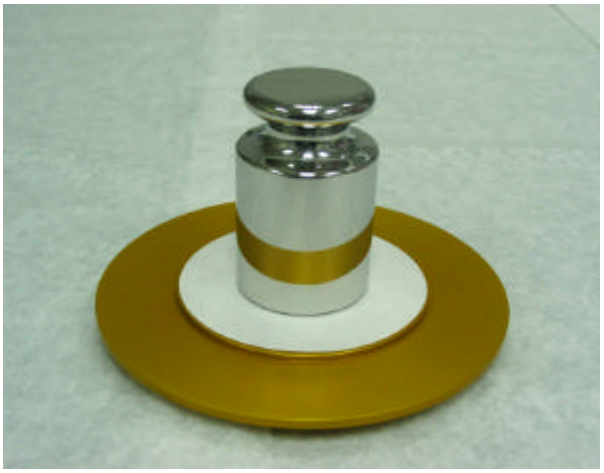


Figure 4 : Malaysian Primary Kilogram

NML-SIRIM maintains a few other pieces of stainless steel kilogram masses which are calibrated at other overseas national measurement standards laboratories traceable to BIPM such as PTB (Germany), NMIJ (Japan) and MSL (New Zealand). These kilogram masses which are also of OIML Class E1 are compared with one another to monitor the stability of the value of the Malaysian Primary Kilogram.

To enable the dissemination of the Malaysian Primary Kilogram to Secondary Standard masses the establishment of a primary standard mass scale from 1 mg to 20 kg is necessary to provide reference values for the calibration of the Secondary Standards. Two sets of weights of OIML Class E1 from 1 mg to 20 kg are used to derive the mass scale through a sequence of build-up and build-down comparisons involving the Primary Kilogram and the two weight sets. Further details of the establishment of the mass scale are described below.

Establishment of the Primary Standards Mass Scale

A mass comparison scheme has been designed for deriving mass values for multiples and sub-multiples of a kilogram, based on the Malaysian Primary Kilogram. Two weight sets, a Reference set and a Control set with nominal values 1 mg to 20 kg are used. The use of two weight sets enables the achievement of better efficiency as compared to possible alternative single weight set comparison schemes. Mass values for weights in the Reference set as well as mass values for the weights in the Control set are derived in this comparison scheme. Values of the Reference set are used as reference values for the dissemination to Secondary Standards while the values for the Control set are used for surveillance purposes.

High precision mass comparators of the following capacities and resolutions are used for the mass comparisons:

Capacity	Resolution	Remarks
5 g	0.1 µg	-
100 g	1 µg	Automated with weight handler
1 kg	1 µg	Automated with weight handler
10 kg	10 µg	Automated with weight handler
20 kg	0.5 mg	Automated with weight handler

A method of calculation involving the use of suitable matrices to determine the mass values and measurement uncertainties from the mass comparison results is employed. An important feature of the method of calculation is the use of the least squares method of analysis which allows ready detection of errors in entered data and poor weighing results through analysis of residuals.

A computer program is specially written to perform the calculation and analysis using Microsoft Excel Workbooks. The program includes all corrections and uncertainty contributions necessary for calculation of mass values to uncertainties better than 1 part in 10^8. The uncertainty of the primary standards mass scale so established meets OIML Class E1 as regards to the uncertainty of measurement.

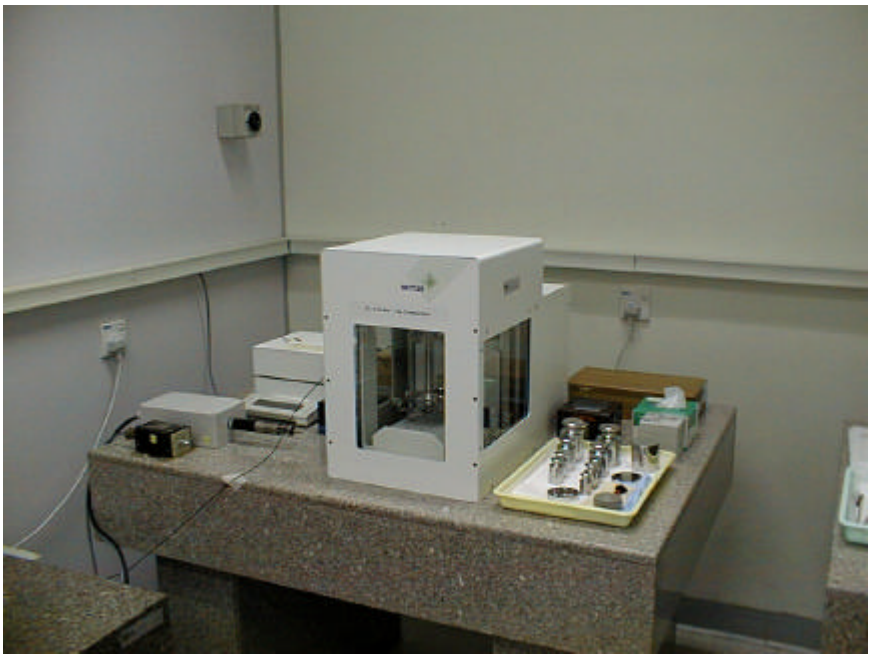


Figure 5 : A 1 kg Mass Comparator used in the mass comparisons from 200g to 1 kg.

National and International Traceability

Figure 6 summarizes the establishment and maintenance of mass standards as prescribed under the Weights and Measures Act 1972.

Through the hierarchy of standards described above measurements for trade and for law enforcement can thus be traced ultimately to the Malaysian Standard Kilogram which is presently traceable to the international prototype kilogram via the national kilograms of overseas national metrology laboratories. It is the intention of Malaysia to arrange for the traceability of the Malaysian Primary Kilogram directly to the kilogram standards maintained by the International Bureau of Weights and Measures (BIPM) in the near future.

Every effort is made to ensure that the highest standards of measurement reliability and accuracy are achieved in the Mass Laboratory in SIRIM Berhad. A system of regular calibration and surveillance of the mass standards and comparators is set up to monitor, and if necessary take corrective action for, both steady and unexpected changes in mass of the Reference and Control Sets used in the establishment of the primary standard mass scale.

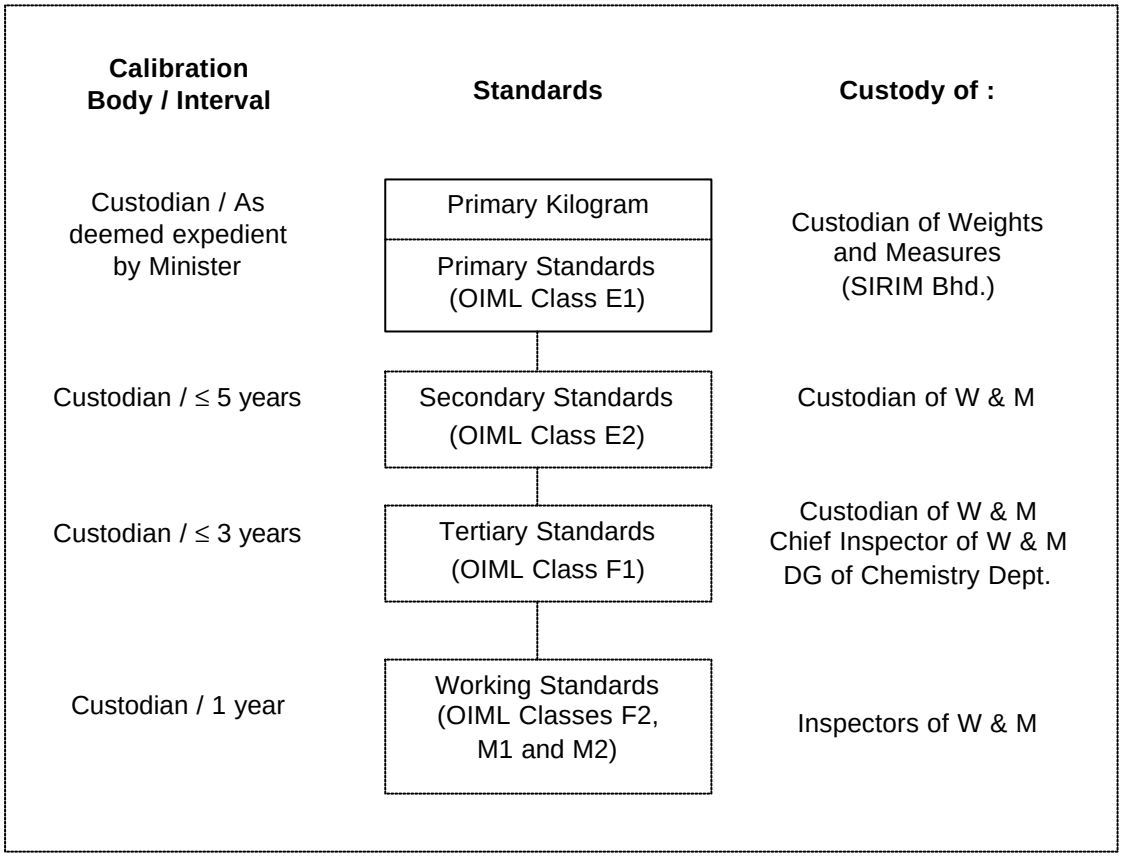


Figure 6 : Establishment and maintenance of Malaysian standards of mass under the Weights and Measures Act 1972

International Harmonization and Mutual Recognition

With the current global developments towards greater promotion of free trade under the WTO Technical Barriers to Trade Agreement there is a need for Malaysia to align and harmonize, and establish mutual recognition arrangement in the field of metrology, both legal and scientific.

Malaysia became a full member of the Metre Convention in September 2001 following which it was accepted as a signatory of the CIPM Mutual Recognition Arrangement (MRA). The CIPM MRA provides for the global recognition of the equivalence of the national standards and measurement capabilities of member states. Such recognition is achieved through a rigorous process of evaluation coordinated by BIPM based very much on the performance of key comparisons of measurement standards organized by regional metrology organizations or BIPM.

In the field of mass metrology Malaysia (represented by NML-SIRIM) has successfully participated in a key comparison of 1 kg mass standards organized by the Asia-Pacific Metrology Programme which was held between July 1999 and March 2001. Two pieces of 1 kg masses were circulated to all participating laboratories. Fifteen national metrology laboratories participated in this intercomparison; CSIRO-NML (Australia), NIS (Egypt), SCL (Hong Kong), NPLI (India), NMIJ (Japan), NML-SIRIM (Malaysia), MSL (New Zealand), ITDI (Philippines), KRISS (Republic of Korea), SPRING (Singapore), CSIR-NML (South Africa), NSCL (Syria), CMS (Taiwan), NIMT (Thailand) and VMI (Vietnam). The results obtained by Malaysia showed a very good degree of equivalence relative to the CCM.M-K1 key comparison reference value. The CCM.M-K1 was organized by BIPM much earlier.

Malaysia will be participating in key comparisons involving the sub-multiples and multiples of the kilogram (APMP.M.M-K2 and APMP.M.M-K5) which are currently being planned by the Asia Pacific Metrology Programme for the near future.

Malaysia's current calibration and measurement capabilities in the field of mass measurements which have been accepted and recognized under Appendix C of the CIPM MRA can be viewed at the BIPM website : http://www.bipm.fr. The capabilities are for services ranging from 1 mg to 20 kg.

Being a member of the Metre Convention now Malaysia is planning to send its Primary Standard Kilogram to BIPM for the next calibration to obtain traceability to the international prototype kilogram. It will however still continue to collaborate closely with other countries in bilateral or multilateral comparison of mass standards. The sending of the Malaysian Primary Kilogram for calibration in BIPM is expected to result in a much lower (better) expanded uncertainty of typically of 0.026 mg (k=2. 95% c.l.) for its calibrated value.

In the area of Legal Metrology, Malaysia has been a corresponding member of the International Organization of Legal Metrology (OIML) since 1989. It is a full member of the Asia Pacific Legal Metrology Forum (APLMF) since its inception in 1994. At the ASEAN level it is a member of the ACCSQ Working Group Legal Metrology. Malaysia's participation in meetings and activities of these organizations has enabled the country to keep abreast with the latest developments in legal metrology and work towards regional and international harmonization and alignment of the country's legal metrology system.

Current Developments in Legal Metrology

Three significant developments currently being undertaken are :

(i) the drafting of a new National Measurement System Act and,

(ii) the privatization of the technical verification services of the weighing instruments presently undertaken by inspectors of weights and measures.

(iii) the construction of a new building for the National Metrology Laboratory, SIRIM Berhad.

Drafting of the National Measurement System Act

The National Measurement System Act seeks to coordinate the national measurement system of the country through prescribing the use of a uniform system of units of measurement based on the SI and streamlining requirements for traceability of measurements made for any legal purpose, including health, safety, forensic and environmental matters. It will also provide for measurement traceability for use in industry and science, and for purposes of international trade. The Act has been drafted in a manner to take into account Malaysia's present and future needs within the context of the international standards and conformance environment. The draft Act is in its final stage of legal process.

Concurrent with the drafting of the above Act the existing Weights and Measures Act 1972 is also being reviewed and a number of amendments on provisions relating to the establishment of national standards has been identified and proposed to avoid overlapping provisions between the two legislations. When completed both Acts will be tabled together for approval by the Parliament.

Privatization of Weights and Measures Services

The privatization of the technical verification services presently carried out by the inspectors of weights and measures involve the taking over of the equipment and assets of the weights and measures offices by a private company. The private company shall be responsible for undertaking the technical verification of weights and weighing and measuring instruments used for trade purposes. To ensure the reliability and technical competence of the private

company undertaking this task it is anticipated that requirements will be imposed for the company to implement and seek accreditation for its laboratories to meet with ISO/IEC 17025.

The Chief Inspector of Weights and Measures, Ministry of Domestic Trade and Consumer Affairs shall however continue to be the regulatory authority on all matters related to weights, measures and instruments for weighing or measuring for use for trade.

The traceability requirements for the Working Standards (and some Tertiary Standards) of mass are however not expected to be changed; they must still be verified and certified against the relevant standards maintained by the Custodian.

New Building for National Metrology Laboratory

A new building with specially designed and environmentally controlled laboratories is being constructed at a new site in Sepang located about 12 km from the Kuala Lumpur International Airport. This new building will house the primary standards and associated precision measuring equipment. The building is expected to be completed in early 2004. The mass laboratories have been upgraded and planned to cater for future expansion in the new building.

Future Direction

In the wake of the strong regulatory reform movement towards a more open and competitive international trade and investment Malaysia hopes to align and harmonize its legal metrology system with OIML recommendations and look forward to the establishment of mutual recognition arrangements with other countries. Malaysia intends to remain active in regional and international legal metrology activities towards achieving these objectives.

Acknowledgements

The author wishes to thank Mr. Mohd Ismail Bin Md Yunus, Director of Legal Metrology, Ministry of Domestic Trade and Consumer Affairs for his valuable assistance in making available samples of old standards of mass used in the 1960s and providing useful information on weights and measures enforcement in Malaysia. His record of appreciation also goes to Mr. Md Nor Bin Md Chik, Senior General Manager, National Metrology Laboratory for his support and comments on the paper.

References

1. J. Bryant, Weights and Measures in Malaysia – A Survey and Recommendations, June 1967.

2. Laws of Malaysia, Act 71, Weights and Measures Act, 1972.

3. Andrew C. Corney, Preparation of a National Measurement Law to provide a Full Legal Basis for a National Measurement System, 21 November 1996, IRL (NZ) Report for SIRIM Bhd.

4. M. T. Clarkson, Establishment of a Mass Scale from 1 mg to 20 kg based on the Malaysian Primary Kilogram, using mass comparators, 15 May 1997, IRL (NZ) Report for SIRIM Bhd.

Measurement Uncertainty and Competence of Verification Organizations

Pao-Chu Kao

Center for Measurement Standards

Industrial Technology Research Institute

321 Kuang-Fu Rd., Sec. 2, Hsinchu 300, Chinese Taipei

Key words: measurement uncertainty, competence, verification organization, capability consistency, error normalized value, maximum permissible errors, legal metrology, target value of measurement uncertainty.

ABSTRACT

The measurement result from a laboratory is not necessary to be used for compliance determination. As mentioned in ISO/IEC 17025 [1], if the result is to be used for compliance description, the measurement uncertainty must be taken into consideration. Because of different measurement capabilities in each laboratory, the evaluated measurement uncertainties are also different. If the same instrument is sent to different laboratories with different measurement uncertainties to execute the calibration, there will be diverse compliance determination from laboratory results. In the field of legal metrology, in order to avoid the influence from the differences in competencies of verification organizations (i.e. the differences of measurement uncertainty) to the compliance determination, suitable tolerances of verification regulations based on the proper capabilities of the verification organizations should be followed. This article briefly introduces the application situation in our country, and suggests a cooperative solution for this existing problem among member countries.

1. INTRODUCTION

According to the ISO GUM [2], measurement uncertainty is a "parameter, associated with the results of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measuerand."

The evaluation of measurement uncertainty cannot only discover improvements in the measurement procedures and determine if the measurement result is compliant with the regulations, but can also exhibit laboratory capabilities. According to the ISO/IEC 17025, the factors effecting the evaluation of measurement uncertainty include staff, facilities, environmental conditions, testing method and method validation, instrument, measurement traceability, and the handling processes for testing and calibration pieces. The contribution to total measurement uncertainty by each factor makes a significant difference to the testing types, in that the measurement uncertainty for each verification organization is different from others due to different effects by the above-mentioned factors. Generally, the smaller the measurement uncertainty, the better the laboratory capability. However, if the organization seeks greater performance, the time and effort involved will also be considerably greater. For the establishment of verification regulations in Taiwan, international regulations are taken into account, but not the items effecting the measurement individually for suitable evaluation. This may result in the tolerances being compliant with the international regulations but not achievable or exceeding the capabilities of the verification organization in our country. Thus, the influences could be at fault in the determination of compliance or wasting resources. Therefore, it is highly important to define the reasonable target values of measurement uncertainty based on requirement.

2. CAPABILITY CONSISTENCY OF VERIFICATION ORGANIZATIONS

2.1 Consistency determination rule

ISO/IEC 17025 may be understood as a normal request for capabilities in testing and calibration laboratories, and is widely accepted by each accreditation organization in the world. According to the regulation request, the assessment of verification organizations' quality system and technical capability is implemented by the domestic authority (Bureau of Standards, Metrology, and Inspection, Ministry of Economic Affairs) based on ISO/IEC 17025. The biggest problem during the implementation in the first few years is how to confirm the consistency of the verification results from each organization. One of the regular missions of the BSMI each year is to execute measurement comparisons between verification organizations. The intercomparison results often need to be transformed into a performance statistic, to aid interpretation and to allow comparison with defined goals. It usually employs the normal determination rules (error normalized, E_n) used in the calibration field for evaluation of

performance. The error normalized (En) with respect to the stated uncertainty can be calculated as the following equation [3,4,5,6]:

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

Where V_{lab} is the participant's result and V_{ref} is the assigned value (reference value is one of the procedures available for the establishment of assigned values and provided by a reference laboratory); U_{lab} is the uncertainty reported by the participating verification body and U_{ref} is the total uncertainty of the reference value (including any allowance for drift or instability of the artifact). Both U_{lab} and U_{ref} must be calculated in a manner consistent with the ISO *Guide to the expression of uncertainty in measurement.* Both uncertainties are at a 95% confidence level. If the absolute value of E_n is less than or equal to one, the performance of the participant is satisfactory; and if the absolute value of E_n is lager than one, the performance of the participant requires further investigation. The participant should provide reasonable explanation or corrective action plan for a unsuccessful performance.

If the verification organization was not requested to evaluate the measurement uncertainty of result, it could not employ the E_n value as the determination rule. The simple difference between the participant's result and assigned value (V_{lab} -V_{ref}), percent difference ($\frac{V_{lab} - V_{ref}}{V_{ref}} \times 100$), percentile or rank, z score ($z = \frac{V_{lab} - V_{ref}}{s}$, where s is an appropriate estimate of variability) are commonly used statistics, only if the assigned value V_{ref} is known. The assigned values are hard to acquire in certain fields, especially for the ones with traceability infrastructure not completely established. Whenever the assigned value is not acceptable, the significance test in the statistics may be used, which uses the measurement values and standard deviations from each verification organization for the F-test and t-test. F-test is used to test the equivalence of variance of two participants and is calculated as:

$F = \frac{s_1^2}{s_2^2}$, Where s_1^2 and s_2^2 represent variance of two participants.

t-test is used to test the equivalence of test result of two participants. If F-test proves the equivalence of s_1^2 and s_2^2, t can be calculated as $t = \frac{\overline{x_1} - \overline{x_2}}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$; if F-test proves unequivalence of s_1^2 and s_2^2, then $t = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$,

Where $\overline{x_1}$ and $\overline{x_2}$ represent test result from two participants,

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

For example, the standard alcohol gas required during verifying evidential breath analyzer cannot be obtained from any gas company or national measurement institute in our country. Thus, E_n and other commonly used statistics for performance evaluation are not suitable, and the direct comparison between the results from two verification organizations is probably the best method to be applied. Sometimes the statistical results will exhibit considerable differences, but it is hard to determine if such differences will effect the determination of verification due to the insufficient information. The significance test for 0.250 mg/l is shown in Table 1 and 0.550 mg/l shown in Table 2. We can conclude that there is statistically significant difference between the two test results.

2.2 Compliance determination

Concerning the determination of exceeding specification for the result by the verification organization, there are various methods to handle the measurement uncertainty. As mentioned in the ISO/IEC 17025, the measurement uncertainty should be taken into consideration for the compliance description. The ordinary determination rule is that, when the measurement result plus the uncertainty is within the specification (or error boundaries), it would be certified, otherwise it would fail when falling outside the specification, with all others falling into the gray area. The basic rules in the document from ILAC --Guidelines on Assessment and Reporting of Compliance With Specification [7] are the same as the above-mentioned, but have further description on the gray area (see Fig. 1). The determination rules in the legal metrology field are slightly different [8] (see Fig. 2). Because it is assumed that the measurement uncertainty is rather small, it could be determined as compliant only if the verification result from the testing instrument is smaller than or equal to the maximum permissible error on verification (MPEV). Nevertheless, if the measurement uncertainty of the verification result is bigger than one third of the MPEV, there is a large probability for fault determination. Especially in criminal cases, if a drunk driver kills someone while driving, whether or not the alcohol content in the body exceeds the legal value will be effected by the accuracy of the evidential breath analyzer, which has great influence upon the defendant's case.

3 DIFFICULTIES DURING EXECUTION

3.1 Traceability

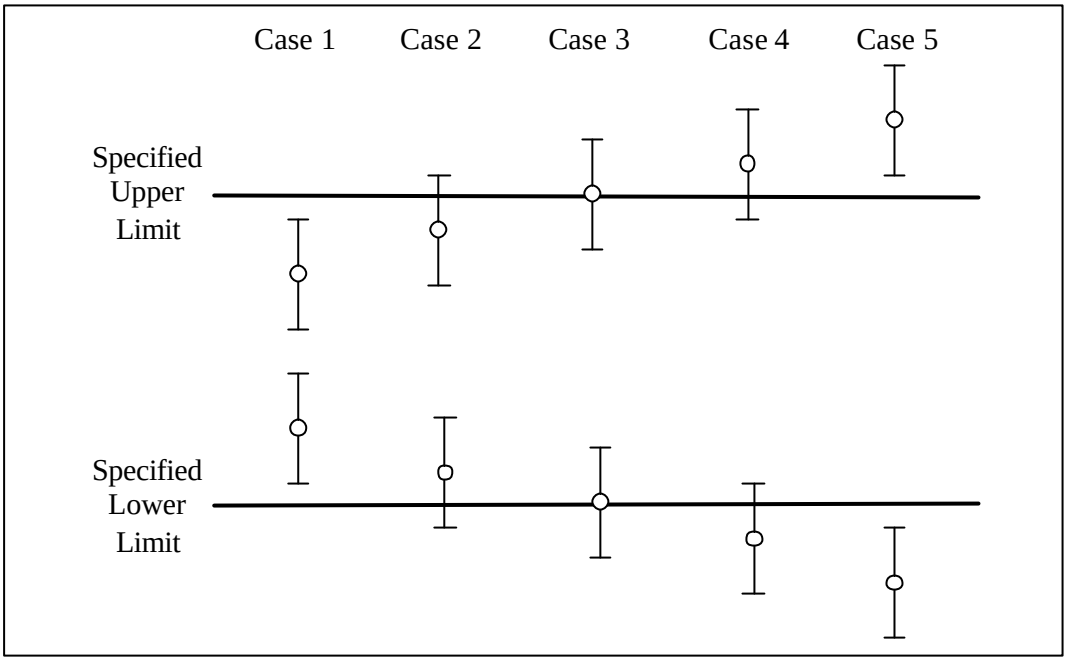


Fig. 1 Uncertainty and Specification Limit

Case 1: Measurement result complies with the specification
Case 2: Changing confidence level to less than 95%, a compliance statement may be possible.
Case 3: Compliance determination depends on the specification limit definition.
Case 4: Changing confidence level to less than 95%, a compliance statement may be possible.
Case 5: Measurement result not complies with the specification

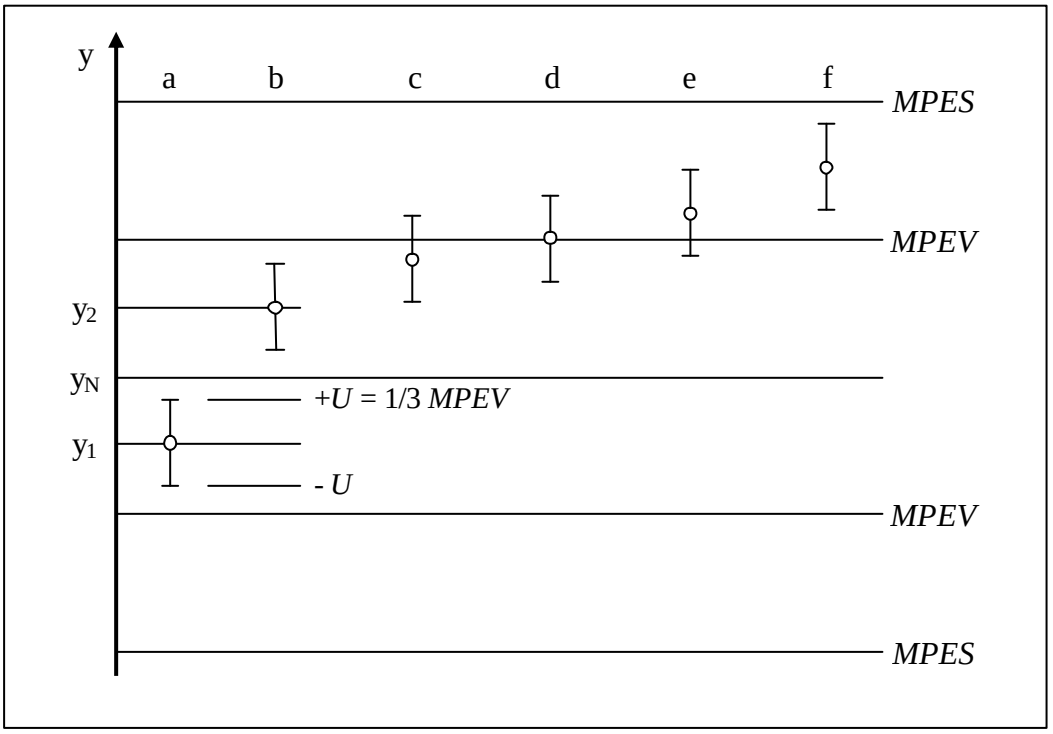


Fig. 2 Influence of the uncertainty of measurement results y_i on conformity assessment in verification

Table 1 Significance test result for 0.250 mg/l

F- test

	1st	2nd
Avg	0.00448	-0.00052
var.	2.09E-06	3.34E-06
# of observation	25	25
df	24	24
F	0.626122	
P(F<=f)one-tailed	0.129253	
Crit.val.0ne-tailed	0.504093	

t -test:
equal variance assumed

	1st	2nd
Avg	0.00448	-0.00052
var.	2.09E-06	3.34E-06
# of observation	25	25
Pooled var.	2.72E-06	
df	48	
t	10.72195	
P(T<=t) one-tailed	1.23E-14	
Crit.val.one-tailed	1.677224	
P(T<=t) two-tailed	2.46E-14	
Crit.val.two-tailed	2.010634	

single factor

set	No	Sum	Avg.	Var.
1st	25	0.112	0.00448	2.0933E-06
2nd	25	-0.013	-0.00052	3.3433E-06

ANOVA

Source of variation	SS	df	MS	F	P-value	Crit. Val.
between groups	0.000313	1	0.0003125	114.960147	2.46E-14	4.042647
within groups	0.00013	48	2.718E-06			
total	0.000443	49				

Table 2 Significance test result for 0.550 mg/l

F test

	1st	2nd
average	0.0096	-0.002
variance	1.83333E-06	3.75E-06
# of observation	25	25
df	24	24
F	0.488888889	
P(F<=f) one-tailed	0.042994881	
critical value: one-tailed	0.504092768	

t -test, assumed unequal variance

	1st	2nd
average	0.0096	-0.002
variance	1.833E-06	3.75E-06
# of observation	25	25
	0	
df	43	
t	24.546027	
P(T<=t) one-tailed	3.181E-27	
critical value: one-tailed	1.6810714	
P(T<=t) two-tailed	6.363E-27	
critical value: two-tailed	2.0166908	

Single factor

set	No	Sum	average	variance
1st	25	0.24	0.0096	1.83333E-06
2nd	25	-0.05	-0.002	0.00000375

ANOVA

source of variation	SS	df	MS	F	P-value	Crit value
between groups	0.001682	1	0.001682	602.5074627	8.068E-29	4.042647
within groups	0.000134	48	2.79167E-06			
total	0.001816	49				

The measurement uncertainty and the measurement traceability are the two essential factors for achieving the consistency request. As mentioned in the ISO/IEC 17025, the result should be traced to the SI, but if not, it may be traced to other suitable standards. It is also acceptable to trace to the certified reference material provided by qualified suppliers in certain technical fields. If the domestic cannot find a suitably traceable channel, we permit the result to trace to foreign standards. Currently, the biggest problem is that the accreditation of manufacturers for reference material is not very prevalent. Therefore, the verification organization cannot ensure the correctness of the measurement uncertainty for the verification results if it could not acquire the reference material that complied with the requirement (such as stability, uniformity, and measurement uncertainty).

3.2 Measurement Uncertainty

The factors should be taken into consideration for the evaluation of the measurement uncertainty including test method, test procedure, instrument and facility, environment, staffs, and even personnel proficiency. In our country, the methods and procedures followed by the certification organizations typically are according to the associated regulations outlined by the authorities (BSMI). Thus, whether the regulations are defined completely and precisely will influence the measurement uncertainty of the verification results. The evaluation of the measurement uncertainty generally is based on the Guide to the Expression of Uncertainty in Measurement, wherein the evaluation steps include:

--Establishing a mathematical model which can express the relationship between the measurand and the input quantities;
--Determining the estimated value of input quantity;
--Evaluating the standard uncertainty u of each input estimate;
--Evaluating the covariance if necessary;
--Calculating the result of the measurement;
--Determining the combined standard uncertainty u_c of the measurement result;
--Supplying an expanded uncertainty U if required; and,
--Reporting the result of the measurement together with it's combined standard uncertainty or expanded uncertainty.

If the verification organizations are requested to proceed with the evaluation of uncertainty according to the above-mentioned rules, the first problem is the establishment of a mathematical model, especially for the system with complicated measuring theories (such as the verification of evidential breath analyzer); then, the acquirement and confirmation of the associated information for each input quantity (for example, the accuracy, repeatability, reproducibility, uncertainty, traceability of reference standard, etc.). These two factors result in the difficulties during execution.

3.3 Problems in evaluating measurement uncertainty

The challenges of evaluating measurement uncertainty in Taiwan are as follows. There are two designated verification bodies carrying out the verification service of evidential breath analyzer. Testing apparatus of these two bodies was provided from different sources (i.e. measuring principle was different, one designed for wet gas and another one for dry gas; both testing apparatus were modified to fit the requirement of our own regulation—wet and dry gas system composed in one device) and resulted in different approaches used for evaluating measurement uncertainty. The first organization established a very simple equation, $Y=X$, and they considered evidential breath analyzer's resolution, pressure transducer, mass flow meter, and standard gas as the error sources. During the evaluation procedure, standard uncertainties were estimated using type B by assuming all error sources with rectangular distribution. Finally the combined standard uncertainty, u_c , was obtained by root sum square the standard uncertainty values of all error factors ($u_c = \sqrt{u_{EBA}^2 + u_p^2 + u_m^2 + u_s^2}$).

The second organization did the work in a different way. Relatively complicated mathematical models were suggested, $Y = C_{dry} = f(C_{std}, C_{reg}, C_{rep}, C_{mea})$ for the dry system and $Y = C_{wet} = f(C_{air}, C_y, C_{ir}, C_p)$ for the wet system, wherein C_{std} ,C_{reg} ,C_{rep} ,and C_{mea} ,represent concentration of dry alcohol gas, linear regression analysis of concentration, repeatability, and stability of plateau respectively; C_{air} ,C_y ,C_{ir} ,and C_p represent concentration of wet alcohol gas, linear regression analysis of concentration, repeatability, and stability of plateau respectively. The combined standard uncertainty, u_c , was derived by the law of propagation of uncertainty. Both the expanded uncertainties of these two organizations were calculated with a 95% confidence level. The measurement uncertainty of the first organization is approximately two to three times the measurement uncertainty of the second one, and usually the measurement uncertainty of these two verification bodies cannot fit the requirement of regulation, at least equal to or less than one third of the maximum permissible error on verification. This makes the compliance determination impossible.

3.4 What can be help?

3.4.1 Guide for measurement uncertainty in legal metrology

If there is an organization which can prepare the writing guidelines of uncertainty evaluation for the execution items of legal metrology, such as EURACHEM/CITAC Guide [9], it could lead the verification organizations to evaluate their own capabilities in a consistent manner, and further confirm if they can achieve the target values of measurement uncertainty (the target values are defined by the authority).

3.4.2 Target value of measurement uncertainty

As for the configuration of target values for measurement uncertainty, it should be requested to be lower than one third of the permitted tolerance, generally. Typically, the permitted tolerances in our country are defined according to international rules. Nevertheless, there are a few points to be noted in the permitted tolerance process:

(1) According to international regulations, instruments have to pass the type approval test before executing the verification, except for those with a simple design that can be tested easily. Due to efficiency, normally the verification may only make one or a few measurements (observations), so that the discrete data for the measurement values from the instrument can be obtained from the type approval. The data is the important source for the evaluation of measurement uncertainty for the following verification results. Because of limited budget and time, some items do not take the type approval, which means they cannot acquire the discrete data for the measurement values from the instrument.

(2) The maximum permissible error on verification (MPEV) and the maximum permissible error on service (MPES) are related to the desired purpose of the instrument, and should be decided by suitable measurement techniques. Currently, our country's definition of the initial verification and subsequent verification tolerances for the verification regulations is synchronous with international practice. However, in order to follow the domestic situation and status, the functions of some legal measurement instruments in our country are different from those in other foreign countries. In certain cases, even verification facilities are different. It is necessary then that the maximum permissible errors on verification and service be self-confirmed under such circumstances.

(3) In the chemical, biological and clinical fields, because of the differences of measurement methods (or procedures), differences will sometimes occur concerning measurement results, even for the measurement uncertainty. Thus, the associated verification regulations for the measurement instruments should be included with complete and detailed standard operation procedures (including the data processing), so as to ensure the consistency of each verification organization following the regulations for job execution. This will reduce the deviation of evaluation for the measurement uncertainty.

4. CONCLUSION

4.1 Suggestions for cooperation among member economies

The uncertainty and measurement traceability are the two essential factors for determining quality. There are many legal measurement items concerning people's livelihood, trading, environmental protection, and traffic safety. If each member economy is to establish the associated techniques themselves, the techniques will be costly and difficult. If member economies can work together, it will help develop a more robust system in each country. Moreover, we suggest the establishment of a task group for measurement uncertainty, to develop a guiding document based on case studies according to the rules of ISO GUM, to provide the users with a consistent evaluation procedure. Furthermore, we suggest holding a regional conference each year, which can provide an experience exchange for technical people engaged in legal measurement. Due to advanced network development, it might be more advantageous to rapid information exchange to set up web sites or on-line forums.

4.2 Further development in Taiwan

A project has been contracted between the Center for Measurement Standards, Industrial Technology Research Institute and the Bureau of Standards, Metrology and Inspection this year. The main purpose of this project is to establish an evaluation model for the consistency of the designated verification organizations. There are three parts to be investigated further:

(1) Is there any other international standard (or guide) that should be referred to for assessing the quality system and technical capability of a designated verification body, except for ISO/IEC 17025?
(2) Since measurement uncertainty is not as popular as in the calibration field, what is the international requirement for legal metrology persons?
(3) How should intercomparisons be executed, especially if reference value is not available?

As pertaining to item one, ISO/IEC 17020 and ISO Guide 65 have been discussed in addition to ISO/IEC 17025, and a comparison report will be produced at the end of the project. In regard to the case of only one body designated for verification work, measurement comparison is not possible; therefore, increasing on-site technical assessment frequency and attending proficiency test held by other proficiency test providers is encouraged.

Concerning item two, measurement uncertainty has not been a requirement for some service fields of legal metrology in the past; therefore, the idea for target value of measurement uncertainty has to be promoted among the verification bodies. Target value of the measurement

uncertainty should be designed according to the maximum permissible error on verification (MPEV) determination based on domestic need [10]. Demonstration of measurement competence should be assessed by technical experts.

A standard operating procedure for executing intercomparison among designated verification organizations has been drafted and discussed in November 2003 by a working group, which is composed of proficiency test experts, responsible staff from the Bureau of Standards, Metrology and Inspection, and representatives from various designated verification bodies. The consensus is that an execution procedure should be in harmony with international standards (such as ISO Guide 43-1, ISO Guide 43-2, APLAC PT001, APLAC PT004…); and if reference value for intercomparison is not available inside Taiwan, participating in proficiency tests held by other countries or suitable providers should be acceptable by the authority. The statistic technique used for performance evaluation still needs investigation. The draft will be finished by the end of this year.

REFERENCES

1. ISO/IEC 17025 (1999) General requirements for the competence of testing and calibration laboratories. ISO.
2. ISO GUM (1995) Guide to the expression of uncertainty in measurement, ISO.
3. APLAC PT 001, Calibration Interlaboratory Comparisons, 2003.
4. APLAC PT 004, Measurement Audits, 2003.
5. ISO GUIDE 43-1, Proficiency Testing by Interlaboratory comparisons—Part 1: Development and Operation of Proficiency Testing Schemes, 1997.
6. ISO GUIDE 43-2, Proficiency Testing by Interlaboratory comparisons—Part 2: Selection and use of proficiency testing schemes by laboratory accreditation bodies, 1997.
7. ILAC Guide 8 (1996) Guidelines on Assessment and Reporting of Compliance with Specification. http://www.ilac.org/
8. Schulz, W.; Sommer, K.-D.: *Uncertainty of measurement and error limits in legal metrology*: OIML Bulletin, October 1999, pp. 5-15.
9. EURACHEM/CITAC (2000) Quantifying uncertainty in analytical measurements. http://www.eurachem.ul.pt/index.htm
10. Paul De Bievre, Philip D. P. Taylor, Klaus Brinkmann, "*Demonstration" vs. "designation" of measurement competence; the need to link accreditation to metrology,* Accred Qual Assur, 7:215-216, 2002.

APEC

35 Heng Mui Keng Terrace Singapore 119616
Tel:(65)6775 6012 Fax:(65)6775 6013
Email:info@mail.apecsec.org.sg
Website:www.apec.org

Inquiries should be directed to:
Secretariat
Asia-Pacific Legal Metrology Forum
NMIJ/AIST Tsukuba Central 3-9
1-1-1 Umezono, Tsukuba, Ibaraki 305-8563, Japan
Tel:(81)29 861 4362 Fax:(81)29 861 4393
Email:e.sec@aplmf.org
Website:www.aplmf.org

APEC#204-CT-01.4 ISBN 4-9901968-0-5