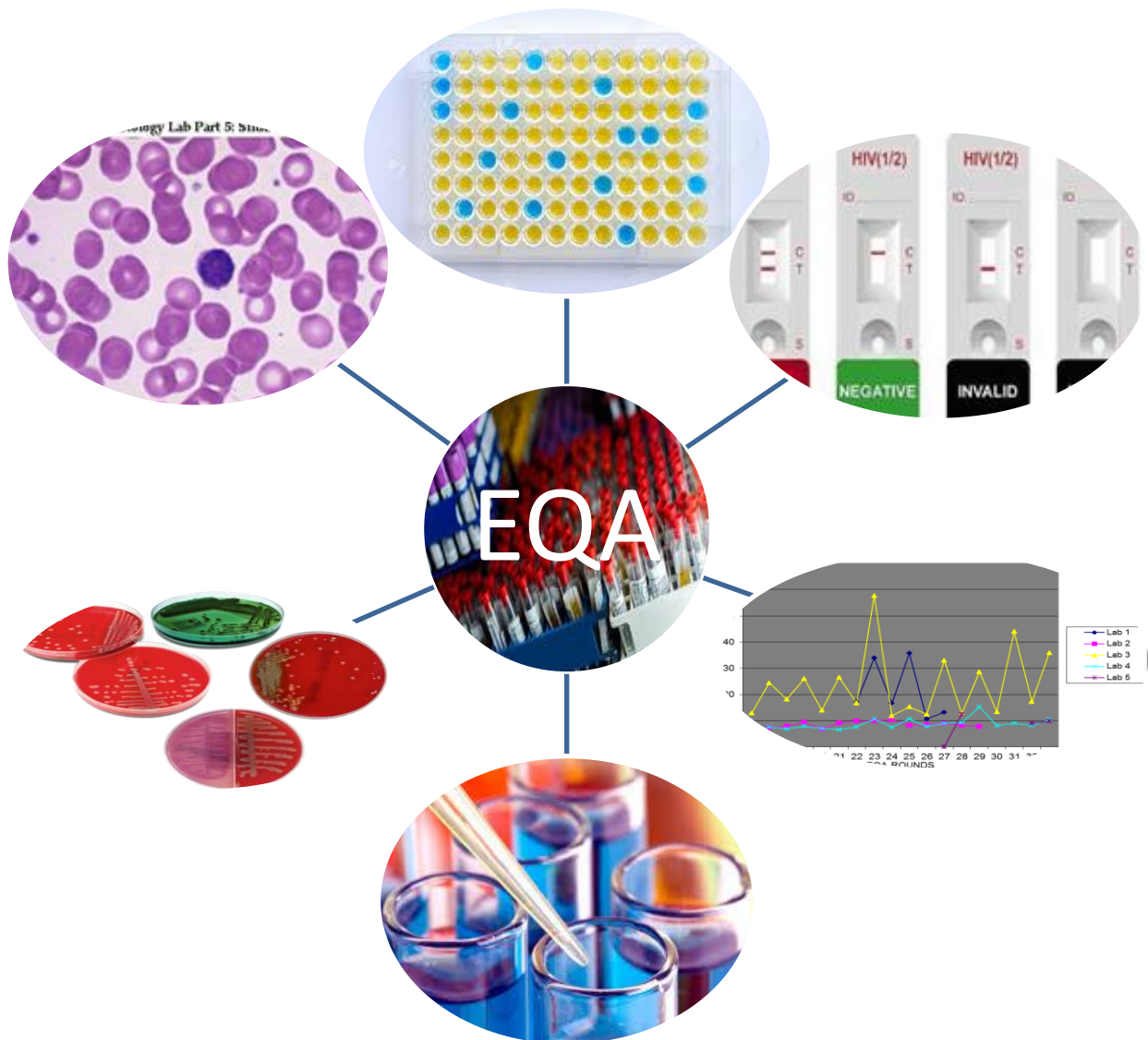




WHO manual for organizing a national external quality assessment programme for health laboratories and other testing sites



WHO PREQUALIFICATION TEAM:
DIAGNOSTICS



World Health
Organization

**WHO manual for organizing a national external
quality assessment programme for health
laboratories and other testing sites**

WHO Library Cataloguing-in-Publication Data

WHO manual for organizing a national external quality assessment programme for health laboratories and other testing sites.

1.Laboratories - standards. 2.Quality Assurance, Health Care. 3.Point-of-Care Systems.
4.Handbooks. I.World Health Organization.

ISBN 978 92 4 154967 7

(NLM classification: QY 25)

© World Health Organization 2016

All rights reserved. Publications of the World Health Organization are available on [the WHO website \(http://www.who.int\)](http://www.who.int) or can be purchased from WHO Press, World Health Organization, 20 Avenue Appia, 1211 Geneva 27, Switzerland (tel.: +41 22 791 3264; fax: +41 22 791 4857; email: bookorders@who.int).

Requests for permission to reproduce or translate WHO publications –whether for sale or for non-commercial distribution– should be addressed to WHO Press through the WHO website (http://www.who.int/about/licensing/copyright_form/index.html).

The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of the World Health Organization concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Dotted and dashed lines on maps represent approximate border lines for which there may not yet be full agreement.

The mention of specific companies or of certain manufacturers' products does not imply that they are endorsed or recommended by the World Health Organization in preference to others of a similar nature that are not mentioned. Errors and omissions excepted, the names of proprietary products are distinguished by initial capital letters.

All reasonable precautions have been taken by the World Health Organization to verify the information contained in this publication. However, the published material is being distributed without warranty of any kind, either expressed or implied. The responsibility for the interpretation and use of the material lies with the reader. In no event shall the World Health Organization be liable for damages arising from its use.

The named authors alone are responsible for the views expressed in this publication.

Printed in France

Contact:

Irena Prat, Essential Medicines and Health Products: WHO, 20 Avenue Appia, 1211 Geneva 27, Switzerland

This document is available on the internet at: www.who.int/diagnostics_laboratory/en/

Contents

1	Introduction	12
2	Strategies for establishing an EQA programme	14
2.1	Situational analysis.....	15
3	Responsibilities of different stakeholders in organizing a national EQA programme	16
3.1	The government.....	16
3.2	The organizing centre(s).....	17
4	Composition of an EQA organizing centre	19
4.1	Guiding principles	19
4.2	Structure	20
4.3	Personnel	20
4.4	Room facilities	21
4.5	Equipment	21
5	Planning and organization of proficiency testing rounds	22
5.1	Planning for establishment of an EQA programme	22
5.2	Planning for implementation of an EQA programme	22
5.3	Packaging.....	24
5.4	Shipment.....	25
5.5	EQA documents	26
5.6	Data entry and evaluation.....	26
5.7	Proficiency testing round reports	27
5.8	Records.....	27
5.9	Archives	28
5.10	Documents for external use	28
5.11	Scientific materials	28
6	Cost estimation	28
6.1	Financial resources.....	28
6.2	Costs of operation.....	28
6.3	Subscription fees	29
7	Proficiency test items.....	29
7.1	Requirements of proficiency test items.....	30
7.2	Collection and testing.....	30
7.3	Preparation of proficiency test items.....	31
7.3.1	<i>Stability during transportation.....</i>	<i>33</i>

8	Broad considerations on the preparation of specific proficiency test items	33
8.1	Haematology	33
8.2	Bacteriology	35
8.3	Parasitology	37
8.4	Clinical chemistry	39
8.5	Serology	40
8.5.1	<i>Serum or plasma</i>	41
8.5.2	<i>Dried tube specimens</i>	42
8.5.3	<i>Characterizing EQA serology materials</i>	44
8.6	Nucleic acid testing	45
8.7	CD4+ T lymphocyte testing specimens	46
8.8	Tissue-based pathology	47
9	Requirements during evaluation of EQA results	47
9.1	Controlled sample	48
9.2	Homogeneity	49
9.3	Stability	50
9.4	Equity	51
9.5	Comparability	51
9.5.1	<i>Robust statistical techniques for estimating the mean and standard deviation</i>	52
9.5.2	<i>Performance scores</i>	53
9.5.3	<i>Measurement uncertainty</i>	54
10	Proficiency testing round report.....	54
10.1	Response to EQA reports	55
11	EQA as an educational tool.....	56
12	Monitoring and evaluating the programme.....	57

Acronyms and abbreviations

ACD	Adenine citrate dextrose
AIDS	Acquired immunodeficiency syndrome
ASLO	Anti-streptolysin O
CPD	Citrate phosphate dextrose
CPDA	Citrate phosphate dextrose adenine
DBS	Dried blood spot (specimen)
DTS	Dried tube specimen
EDTA	Ethylene diamine tetra acetic acid
EIA	Enzyme immunoassay
ELISA	Enzyme-linked immunosorbent assay
EN	European norm
EQA	External quality assessment
FFPE	Formalin-fixed paraffin embedded
HBsAg	Hepatitis B surface antigen
HIV	Human immunodeficiency virus
IEC	International Electrotechnical Commission
INR	International normalized ratio
IUPAC	International Union of Pure and Applied Chemistry
ISO	International Organization for Standardization
IQC	Process control as per ISO15189
MOH	Ministry of health
MRSA	Methicillin resistant <i>Staphylococcus aureus</i>
NAT	Nucleic acid testing
<i>nIQR</i>	Normalized inter-quartile range
PBS	Phosphate buffered saline
PT	Proficiency testing
POC	Point of care
QC	Quality control
RDT	Rapid diagnostic test
R&D	Research and development
SD	Standard deviation
TPHA	<i>Treponema pallidum</i> heamagglutination
TPPA	<i>Treponema pallidum</i> particle agglutination
UK NEQAS	United Kingdom National External Quality Assessment Scheme
VDRL	Venereal disease research laboratory test
VIM	International vocabulary in metrology
WHO	World Health Organization

Glossary

Definitions in the glossary are from the *International vocabulary of metrology (VIM) (1)*, the International Organization for Standardization (ISO) *Guide on vocabulary and symbols (2)*, and the International Union of Pure and Applied Chemistry (IUPAC) *Compendium of analytical nomenclature (3)*. Highest authority is given to the VIM definitions.

Accuracy (measurement accuracy) closeness of the agreement between a measured quantity value and a true quantity value of the measurand (VIM 2.13:2008).

Assigned value attributed to a particular property of a proficiency item (ISO/IEC 17043:2010).

Calibrator

measurement standard used in calibration (VIM 5.12:2008).

NOTE The term "calibrator" is only used in certain fields.

Consensus value

value derived from collection of results in an inter-laboratory comparison (ISO 13528:2015).

Coordinator

one or more individuals with the responsibility for organizing and managing all activities involved in the operation of a proficiency testing scheme (ISO/IEC 17043:2010).

Imprecision (imprecision of measurement)

the random dispersion of a set of replicate measurements and/or values expressed quantitatively by a statistic, such as standard deviation or coefficient of variation (the presence of random error, variability, or inconsistency) (ISO 3534-1:2006).

NOTE 1 It is defined in terms of repeatability and reproducibility.

NOTE 2 The words "imprecision" and "precision" are often inappropriately interchanged.

Interlaboratory comparison

organization, performance and evaluation of measurements or tests on the same or similar items by two or more laboratories in accordance with predetermined conditions (ISO/IEC 17043:2010).

Limit of detection

result of a measurement by a given measurement procedure for which the probability of an analytically false-negative result is B, given the probability of A of an analytically false-positive result (ISO 17025:2005).

Measurement bias (bias)

estimate of a systematic measurement error (VIM 2.18:2008).

Measurement error: Measured quantity value minus a reference quantity value (VIM 2.16:2008).

Measurement precision (precision)

closeness of agreement between indications or measured quantity values obtained by replicate measurements on the same or similar objects under specified conditions (VIM 2.15:2008).

NOTE 1 The measure of precision is usually expressed numerically by measures of imprecision, such as standard deviation, variance, or coefficient of variation under the specified conditions of measurements.

NOTE 2 The “specified conditions” can be, for example, repeatability conditions of measurements, intermediate precision conditions of measurement, or reproducibility conditions of measurements (see ISO 5725-5:1998/Cor 1:2005).

NOTE 3 Measurement precision is used to define measurement repeatability, intermediate measurement precision, and measurement reproducibility.

NOTE 4 Sometimes “measurement precision” is erroneously used to mean measurement accuracy.

Measurement reproducibility

measurement of precision under reproducibility conditions of measurements (VIM 2.25:2008).

NOTE Relevant statistical terms are given in ISO 5725-5:1998/Cor 1:2005.

Measurement uncertainty (uncertainty of measurement, uncertainty)

non-negative parameter characterizing the dispersion of the quantity values being attributed to a measurand, based on the information used (VIM 2.26:2008).

NOTE 1 Measurement uncertainty includes components arising from systematic effects, such as components associated with corrections and the assigned quantity values of measurement standards, as well as the definitional uncertainty. Sometimes estimated systematic effects are not corrected for but, instead, associated measurement uncertainty components are incorporated.

NOTE 2 The parameter may be, for example, a standard deviation called standard measurement uncertainty (or a specified multiple of it), or the half-width of an interval, having a stated coverage probability.

NOTE 3 Measurement uncertainty comprises, in general, many components. Some of these may be evaluated by Type A evaluation of measurement uncertainty from the statistical distribution of the quantity values from series of measurements and can be characterized by standard deviations. The other components, which may be evaluated by Type B evaluation of measurement uncertainty, can also be characterized by standard deviations, evaluated from probability density functions based on experience or other information.

NOTE 4 In general, for a given set of information, it is understood that the measurement uncertainty is associated with a stated quantity value attributed to the measurand. A modification of this value results in a modification of the associated uncertainty.

Participant: laboratory, organization or individual that receives proficiency test items and submits results for review by the proficiency testing provider (ISO/IEC 17043:2010).

Proficiency testing: evaluation of participant performance against pre-established criteria by means of interlaboratory comparisons (ISO/IEC 17043:2010).

NOTE 1 For the purposes of this International Standard, the term proficiency testing is taken in its widest sense and includes but is not limited to:

- quantitative schemes;
- qualitative schemes — where laboratories are required to identify a component of a proficiency test item;
- single item testing — where one item is sent to a number of laboratories sequentially and returned to the organizer at intervals;
- bulk interlaboratory testing — where laboratories are provided with aliquots or portions of a large homogeneous bulk material, and make the necessary tests and measurements within a defined time period;
- single occasion exercises — where laboratories are provided with a proficiency test item on a single occasion;
- continuous schemes — where laboratories are provided with proficiency test items at regular intervals on a continuing basis;
- sampling — where individuals or organizations are required to take specimens for subsequent analysis; and
- data transformation and interpretation — where laboratories are furnished with sets of data or other information and process the information to provide an interpretation (or other outcome).

NOTE 2 Some providers of proficiency testing in the medical area use the term “external quality assessment”.

Proficiency test item: specimen, product, artefact, reference material, piece of equipment, measurement standard or data set provided to one or more participants, or submitted by participants, in a proficiency testing round (ISO/IEC 17043:2010).

Proficiency testing provider: organization which takes responsibility for all tasks in the development and operation of a proficiency testing scheme. (ISO/IEC 17043:2010)

Proficiency testing round: single complete sequence of distribution of proficiency test items, and the evaluation and reporting of results to all participants in a proficiency testing scheme (ISO/IEC 17043:2010).

Proficiency testing scheme: proficiency testing designed and operated in one or more rounds for a specified area of testing, measurement, calibration or inspection (ISO/IEC 17043:2010).

Quality management system: a set of policies, processes and procedures required for planning and execution (production/development/service) in the core business area of an organization (i.e. areas that can impact the organization's ability to meet customer requirements) (ISO 9001:2015).

NOTE 1 The term “quality management system”, as described in *Quality management systems – requirements* (ISO9001:2015), can be applied to the full spectrum of quality management system needs in the medical laboratory. Similar principles are described in *Medical laboratories requirements for quality and competence* (ISO15189:2012) and in *Conformity assessment – requirement for proficiency testing* (ISO/IEC 17043:2010).

NOTE 2 a system that controls, assures and manages the quality of testing services and ensures that quality objectives are met.

Reference material: Material, sufficiently homogenous and stable with respect to one or more specified properties, which has been established to be fit for its intended use in a measurement process (ISO 13528:2015).

Reproducibility condition of measurement: condition of measurement, out of a set of conditions that includes different locations, operators, measuring systems, and replicate measurements on the same or similar objects (VIM 2.24:2008).

NOTE 1 The different measuring systems may use different measurement procedures.

NOTE 2 A specification should give the conditions changed and unchanged to the extent practical.

Robust statistical methods: statistical methods insensitive to small departures from underlying assumptions surrounding an underlying probabilistic model (ISO/IEC 17043:2010).

Standard (measurement standard): realization of the definition of a given quantity with stated quantity value and associated measurement uncertainty, used as a reference (VIM 5.1:2008).

EXAMPLE 1 1 kg mass measurement standard with an associated standard measurement uncertainty of 3 µg.

EXAMPLE 2 100 Ω measurement standard resistor with an associated standard measurement uncertainty of 1 µΩ.

EXAMPLE 3 Caesium frequency standard with a relative standard measurement uncertainty of 2×10^{-15} .

EXAMPLE 4 Standard buffer solution with a pH of 7.072 with an associated standard measurement uncertainty of 0.006.

EXAMPLE 5 Set of reference solutions of cortisol in human serum having a certified quantity value with measurement uncertainty for each solution.

EXAMPLE 6 Reference material providing quantity values with measurement uncertainties for the mass concentration of each of ten different proteins.

NOTE 1 A “realization of the definition of a given quantity” can be provided by a measuring system, a material measure, or a reference material.

NOTE 2 A measurement standard is frequently used as a reference in establishing measured quantity values and associated measurement uncertainties for other quantities of the same kind, thereby establishing metrological traceability through calibration of other measurement standards, measuring instruments, or measuring systems.

NOTE 3 The term “realization” is used here in the most general meaning. It denotes three procedures of “realization”. The first one consists in the physical realization of the measurement unit from its definition and is realization *sensu stricto*. The second, termed “reproduction”, consists not in realizing the measurement unit from its definition but in setting up a highly reproducible measurement standard based on a physical phenomenon, as it happens, e.g. in case of use of frequency-stabilized lasers to establish a measurement standard for the metre, of the Josephson effect for the volt or of the quantum Hall effect for the ohm. The third procedure consists in adopting a material measure as a measurement standard. It occurs in the case of the measurement standard of 1 kg.

NOTE 4 A standard measurement uncertainty associated with a measurement standard is always a component of the combined standard measurement uncertainty (see GUM:1995, 2.3.4) in a measurement result obtained using the measurement standard. Frequently, this component is small compared with other components of the combined standard measurement uncertainty.

NOTE 5 Quantity value and measurement uncertainty should be determined at the time when the measurement standard is used.

NOTE 6 Several quantities of the same kind or of different kinds may be realized in one device which is commonly also called a measurement standard.

NOTE 7 The word “embodiment” is sometimes used in the English language instead of “realization”.

NOTE 8 In science and technology, the English word “standard” is used with at least two different meanings: as a specification, technical recommendation, or similar normative document (in French « norme ») and as a measurement standard (in French « étalon »). This vocabulary is concerned solely with the second meaning.

NOTE 9 The term “measurement standard” is sometimes used to denote other metrological tools, e.g. “software measurement standard” (see ISO 5436-2: 2012).

Standard deviation for proficiency testing: measure of dispersion used in the evaluation of results of proficiency testing. (ISO/IEC 17043:2010).

Value (value of a quantity value): number and reference together expressing magnitude of a quantity (VIM 1.19: 2008).

EXAMPLE 1 Length of a given rod: 5.34 m or 534 cm.

EXAMPLE 2 Mass of a given body: 0.152 kg or 152 g.

EXAMPLE 3 Curvature of a given arc: 112 m^{-1} .

EXAMPLE 4 Celsius temperature of a given specimen: $-5 \text{ }^\circ\text{C}$.

EXAMPLE 5 Electric impedance of a given circuit element at a given frequency, where j is the imaginary unit: $(7 + 3j) \Omega$.

EXAMPLE 6 Refractive index of a given sample of glass: 1.32.

EXAMPLE 7 Rockwell C hardness of a given sample: 43.5 HRC.

EXAMPLE 8 Mass fraction of cadmium in a given sample of copper: $3 \mu\text{g}/\text{kg}$ or 3×10^{-9}

EXAMPLE 9 Molality of Pb^{2+} in a given sample of water: $1.76 \mu\text{mol}/\text{kg}$.

EXAMPLE 10 Arbitrary amount-of-substance concentration of lutropin in a given sample of human blood plasma (WHO International Standard 80/552 used as a calibrator): 5.0 IU/l, where “IU” stands for “WHO International Unit”.

NOTE 1 According to the type of reference, a quantity value is either:

- a product of a number and a measurement unit (the measurement unit one is generally not indicated for quantities of dimension one) or
- a number and a reference to a measurement procedure, or
- a number and a reference material.

NOTE 2 The number can be complex.

NOTE 3 A quantity value can be presented in more than one way.

NOTE 4 In case of a vector or tensor quantities, each component has a quantity value. EXAMPLE Force acting on a given particle, e.g. in Cartesian components $(F_x; F_y; F_z) = (-31.5; 43.2; 17.0)$ N.

Z score: standardized measure of performance, calculated using the participant result, assigned value and the standard deviation for proficiency assessment.

NOTE 1 A common variation on the z score, sometimes denoted z' (commonly pronounced z-prime), is formed by combining the uncertainty of the assigned value with the standard deviation for proficiency assessment before calculating the z score (ISO 13528:2015).

1 Introduction

Results of clinical testing obtained from laboratories or testing sites at or near the point of care must be as accurate as possible as they have a direct impact on care and treatment, prevention, and control of diseases. Testing consists of pre-examination, examination, and post-examination processes, which require strict implementation of a quality management system (4). The components of this quality system include internal quality (process) control, proficiency testing (PT), and quality improvement. It is only when these components are implemented together that quality improvement or the highest attainable quality of testing can be achieved and consequently can improve health outcomes in terms of disease prevention and control, care and treatment.

The term “external quality assessment” (EQA) is used to describe a method/process that allows testing conducted by a laboratory, testing site or individual user to be compared to that of a source outside the laboratory – of a peer group of laboratories or a reference laboratory or testing site. Successful participation of a testing site in an EQA programme provides objective evidence of the testing service’s competence for customers, accrediting bodies and regulatory agencies, and serves as a unique source of information that is not easily obtainable in other ways (5,6). Importantly, participation in an EQA programme allows for a “peer-review” process towards solving technical and methodological problems to improve the quality of service for each individual testing site as well as to achieve comparability of results among different testing services. For accrediting bodies and regulatory agencies, EQA provides objective data on the quality of delivered services, and has been shown to reflect the quality of testing of patients’ specimens (7,8).

It is important that an EQA programme organizer provides a supportive and enabling programme to help participating sites improve their performance, rather than institutes punitive measures for poor performers. If national regulations require participation of a testing site for licensing and/or accreditation, the EQA provider should ensure that EQA participants give consent before sharing results with relevant authorities. Falsification of data by an EQA participant should be discouraged and the benefits of an honest appraisal of performance as a guide to taking correcting action emphasized. The costs of setting up and running an EQA programme must be viewed against the resources wasted in producing inaccurate testing results and the improvement in patient care which can be achieved by enhancing and maintaining the quality of testing results.

An EQA programme may be organized on a sub-national, national, regional or international basis; each has its advantages and disadvantages. Sub-national and national programmes can prove more responsive to the testing sites and are thus able to provide rapid feedback of results and support to participating testing services; however, they may not be sufficiently robust in statistical analysis, or provide economies of scale to make them cost-effective. In addition, these programmes may not gather data from sufficient methods to provide a “state of the art” evaluation of methods.

National EQA programmes are effective for most commonly performed assays and are very useful in providing data on the accuracy of testing results for national guidelines. On the other hand, international programmes are essential for rare or specialized assays, for which there may not be sufficient sites in one country to provide enough data for robust statistical analysis, or for which particular expertise is required to produce proficiency test items. In addition, regional and international programmes may prove useful in supporting and assessing the performance of laboratories constituting the national EQA programme. Other factors to be considered in the organization of an EQA programme include language barriers, the stability of proficiency test items during transit, especially in extreme environments, and shipment logistics and costs.

In most developed countries, EQA programmes are well established and have contributed greatly to improving the quality of care at all levels of the health care system (7–10). In order to improve the situation in developing countries, governments need to recognize the importance of: implementing quality measures in patient care and disease prevention and control, advocating quality testing at all levels, and allocating adequate resources to ensure quality testing. This will require establishing a national quality testing programme, increasing local expertise to run the programmes, and, where possible, encouraging the local preparation of proficiency test items (11). The implementation of a national quality assurance policy requires a commitment from government, professional societies, laboratory/testing site workers and implementing partners. It is important to recognize that a policy towards improving health care should include EQA programmes for all testing sites (health laboratories and other testing centres); poor testing results have a significant financial, health and social impact and therefore investments set aside to improve quality testing will yield future benefits.

The organization of an EQA programme is a technical process that includes a number of interacting elements, implementation of all of which will be accomplished only after the final stage of organization has been reached. Therefore the implementer should not be deterred by the size and scope of the task, but should rather start on a small scale with achievable objectives. This manual is therefore aimed at stimulating national authorities to recognize the importance of an EQA programme, establish the process of setting up a national programme (if necessary in collaboration with existing providers) and support the testing network involved in running a national EQA programme. The building of national capacity in the local preparation of proficiency test items wherever possible (see Section 8), and the provision of support to testing services that perform poorly, is encouraged. While the general principles for preparing the most common proficiency test items are covered in Section 8, this is an area that may require specialist training and research and development. Specific guidelines and literature should be consulted, where applicable; the development of locally prepared materials may need to be introduced gradually. The fundamental principles and requirements for running an EQA programme are described in detail in ISO/IEC 17043:2010 (12); the reader is advised to consult this standard guidance for more details.

This manual describes some of the strategic, managerial, financial, technical and scientific aspects to be considered in establishing a national EQA programme for clinical laboratories and

other testing services at all health care levels. The manual has been prepared by the World Health Organization (WHO) and partners to fill perceived gaps, and the scope is limited to EQA programmes that deliver proficiency testing. However, it is important to note that other EQA activities as well as internal quality control (process control) measures and other quality elements are an integral part of the quality laboratory management process as defined in ISO 15189:2013 (13).

The scope of the manual is all WHO Member States, with a focus on developing countries as the principal target audience. The intended audience includes ministries of health, programme managers, laboratory managers, testing personnel and other implementing partners and EQA providers.

2 Strategies for establishing an EQA programme

Decision makers need to consider a variety of factors before deciding on development of a national EQA programme, or working with an international programme, or creating a hybrid strategy. Local programmes will have start-up costs, and expertise may not be immediately available. On the other hand, local programmes can address local concerns more effectively and in the long run will have less complex transport issues and lower costs. International programmes are already established but are not necessarily focused on local or regional issues or complexities and can also be costly. Increasingly, transport of biological materials across international boundaries is a concern; the longer transport takes, the higher the likelihood of deterioration of proficiency testing (PT) specimens. Ultimately the decision taken has to be in the best interests of health care and quality of care, and the most effective balance of costs and services.

There are two main strategies that can be used to establish a national EQA programme depending on the local situation. In the first strategy, different nongovernmental organizations provide EQA services that may be competitive or complementary to each other. In the second strategy, a relevant national agency that may include a government ministry or other professional or academic institution with a long-standing interest or passion in supporting laboratory quality, is mandated by the government as an organizing centre to run the programme. These two strategies are not mutually exclusive since some countries combine not-for-profit providers with a government EQA programme. The organizing centre should strive to cover public and private diagnostic service providers to ensure a common standard for performance assessment.

The EQA organizing centre should aim to improve the performance of testing sites, and should not be influenced by conflicting interests or be under any undue influence that may affect the running of the EQA programme. The final goal is for the organizing centre to be accredited to the international standard for providers of EQA programmes (ISO/IEC 17043:2010) (12). All public and private laboratories should be encouraged to participate but, wherever possible, participation should be mandatory.

In large countries, a hierarchical structure, with delegation of responsibilities and supervision to provincial/regional and district laboratories, may be more appropriate for an EQA programme. Responsibilities include distributing the proficiency test items provided by the central organizing centre, supervising peripheral laboratories, education and training, and supporting corrective action at the testing sites. The responsibilities of the organizing centre are described in more detail in Section 3.2.

During the initial phase of implementing an EQA programme, it is preferable to include testing services in the vicinity of the organizing centre and those with a high daily workload as a learning curve to enable gaining of experience in organizing the EQA programme. It is preferable if analytes that have the most critical impact on patients' management or epidemiological surveillance are given priority. After addressing the bottlenecks, and as experience is gained, the EQA programme can be gradually extended to peripheral testing sites and more analytes can be added.

Proficiency test items may be commercially acquired or locally prepared. The choice of materials should be guided by availability, as well as by funding, human resources, the expertise required to prepare the items locally, and the number of participating testing sites. Recently there have been successful efforts to prepare proficiency test items locally in order to cater for increasing numbers of testing sites, especially at peripheral levels, and to be able to address local pathology. A good example is the use of dried tube specimens (DTS) for HIV and syphilis serology, or quantitative HIV nucleic acid testing (NAT) (for HIV viral load) introduced by the US Centers for Disease Control (14). This approach has proved to be very successful and is widely accepted and used in many developing countries (15,16). However, in areas such as flow cytometry (e.g. CD4⁺ T lymphocyte enumeration), the preparation of suitable material using fixative has proved to be very challenging. Importation of EQA materials can be the preferred option in this situation; however, the high cost, stability of proficiency test items, and customs clearance need to be considered.

Communication at different levels is the basis for the success of an EQA programme. The EQA organizing centre maintains communication with all relevant groups including national and international experts, professional societies for the elaboration of protocols for EQA programmes, ministry-level health authorities responsible for laboratory and testing services, and companies providing laboratory materials and reagents.

2.1 Situational analysis

A risk-based situational analysis of laboratories and other testing services is the basis for planning and implementing an effective strategy for EQA. The assessment should include, but not be limited to, the following aspects:

- number of laboratories or testing centres in the national laboratory network and other testing sites, including their location, availability of transportation, courier or postal services;

- type of laboratory/testing sites (clinical or public health) and their location (zonal regional/provincial, district, primary, community) – it may be necessary to define their role in offering clinical services and providing epidemiological surveillance of diseases;
- essential infrastructure in the laboratories or point-of-care (POC) testing sites (power supply, water supply, distilled water supply, drainage, waste disposal);
- principal investigation(s) and procedures of measurement(s);
- type of laboratory analyses carried out in the various disciplines (i.e. immunology, clinical chemistry, haematology, bacteriology, parasitology, serology, food and water analysis), number of tests performed per year, and laboratory techniques and frequency of their use;
- existence of implementation of laboratory quality management measures;
- availability of support for supervisory and mentoring networks;
- extent and effectiveness of internal quality (process) control procedures;
- inventory, state of performance and maintenance of laboratory equipment;
- calibrators, proficiency test items and their supply;
- number of technical and medical professionals working in the testing centres and their levels of training, experience and qualifications;
- existing infrastructure for transport of proficiency test items from the central EQA organizing centre to the laboratories in a timely and controlled manner;
- financial aspects for provision of the testing service;
- supply chain management system;
- national regulatory and licensing requirements;
- existing national disease programmes;
- existing disease-specific EQA programmes;
- ability to participate in international external EQA programmes;
- existing information technology networks including internet and mobile communication.

3 Responsibilities of different stakeholders in organizing a national EQA programme

3.1 The government

The government, through its responsible ministries and institution(s), supports the running of the national EQA programme with appropriate legislation or regulation(s), and encourages provision of appropriate services taking into account the national priorities.

The main roles of the national authority include:

- Establishing national regulations, guidelines and policies for setting up, running and coordinating one or more EQA programmes that comply with ISO/IEC 17043:2010 (12) and other appropriate international standards.
- Advocating the role of the national EQA programme as an essential component of a broader national laboratory quality management system and quality improvement strategy.

- Identifying target pathology for the EQA programme depending on priorities for clinical and public health laboratory services.
- Guaranteeing financial and logistical support for the programme(s).
- Mobilizing professional organizations and other stakeholders in support of the programme.
- Identifying the appropriate organizing centre(s) for the programme.
- Ensuring the availability of essential and necessary laboratory infrastructure for the organizing centre(s).
- Ensuring that public and private laboratories participate in the programme.
- Strengthening other quality management system elements in all testing centres based on various approaches including the stepwise approach, with ultimate achievement of ISO 15189:2013 certification (13).

3.2 The organizing centre(s)

The organizing centre(s) should have the required competencies to run an EQA programme, supported by relevant experts depending on the type of analytes included in the programme, as described in ISO/IEC 17043:2010 (12). The selection or approval of an EQA organizer and support facilities for the programme is an important process that will inevitably affect the quality, scope and subsequent success of the programme. While the organizing centre will have to address commercial, financial, and organizational concerns, it should be structured so that decision-making can be fair and independent of influence or bias. The organizing centre is expected to establish communication with technical experts and participants in the EQA programme and to be responsible for the organization, appropriate management of PT rounds and timely evaluation of reports from the participants. Every country/region should strive to develop an EQA programme(s) that can ultimately meet international standards.

The main responsibilities of an organizing centre are to:

- Ensure that it has competent managerial and technical personnel with the required authority to implement a successful EQA programme, including an expert advisory committee.
- Determine the minimum level of qualifications and experience needed to run the EQA programme, allocate specific tasks, maintain all relevant educational and competence records of staff.
- Define the objectives of the EQA programme.
- Ensure the availability of adequate equipment and infrastructure to run the programme.
- Plan a sufficient number of PT rounds during a given period depending on the type of EQA programme and prevailing conditions.
- Select appropriate homogeneous and stable proficiency test items for distribution; the quality of the proficiency test items should be known prior to their use, and it is the responsibility of the organizing centre to control their appropriateness as stipulated in ISO/IEC 17043:2010 (12).
- Define the number of proficiency test items to be provided for each PT round.

- Identify an appropriate competent courier service or distribution mechanism for transporting specimens under conditions that ensure public safety, prevent deterioration of specimens, and are accompanied by clear instructions to participants.
- Provide a system by which participants can return PT results for evaluation, and define the time frame for return of results.
- Establish a method for determining the assigned value, which may be a measurement procedure of highest metrological order, a reference procedure, a specified method-dependent procedure, or be the mean or median of all measurements, as specified in ISO 13528:2015 (17).
- Define the measurement uncertainty that is fit for the purpose (17).
- Develop appropriate statistical design and statistical analysis tools to assess all responses from participants within the established time frame, depending on the objectives of the EQA programme.
- Ensure that individual results of assessments remain confidential, and that the confidentiality of individual participants is preserved on presentation of results of a PT round.
- Identify laboratories/testing sites/individual testers performing poorly or unsatisfactorily, and ensure appropriate assistance is provided. Remedial action may include technical advice and support to the laboratory/testing site by designated supervisors, arrangement of workshops and refresher training courses, and distribution of additional testing materials as appropriate. The organizing centre may also make recommendations to the central laboratory administration to provide appropriate staff as well as essential supplies, equipment or other services to the testing centres to address the weaknesses identified during the PT round.
- Participate in accredited international EQA programmes to ensure performance is up to the required standards.
- Establish a performance monitoring system including national scientific and professional societies that has ultimate responsibility for surveying performance.

A well-staffed and financed organizing centre is critical for establishing and running a successful EQA programme, which supports participants

It is important that EQA participants confirm their commitment, preferably in writing, to the terms and conditions of participation in the programme, including sharing of results with ministry of health coordinating centres and licensing or accreditation bodies, depending on national regulations. This commitment should be made known to the central laboratory or ministry of health (MOH) administration. Under the terms and conditions for participation in an

EQA programme, it should be mandatory that a laboratory takes part in each PT round or informs the organizers of the reason(s) for not doing so. For the programme to be educational, and to help in improving patient care, it is important that participants handle the proficiency test items for measurement in the same way as they handle the specimens they test routinely, as far as possible. Participants must avoid any attempt to manipulate the results or test the proficiency test items separately; items should never be shared with other laboratories, nor must other laboratories ever be asked to perform the testing. Participants should sign the reports to indicate that the investigations were carried out in their own laboratory, using the method(s) indicated in the report. To reduce the chances of miscalculation of results, the PT results should be reported together with the conditions of measurement (e.g. at a given temperature), rather than the participants recalculating the results to meet a specific condition, which may not be routine practice. Results should be reported in the units of measurement required by the programme to ensure valid analysis.

Participants should handle proficiency test items for measurement in the same way as they handle specimens that are routinely tested in their laboratory or testing site

4 Composition of an EQA organizing centre

This section describes a model structure for an organizing centre responsible for running an EQA programme in one or more disciplines. The centre may also function as an independent health care, academic or commercial institution under national regulations. The information contained in this section serves as guidance rather than as recommendations for planning of an organizing centre. In some countries, there may be more than one organizing centre, which may address different speciality areas or levels of implementation of the programmes, coordinated through a central body or national organization.

In practice, the local situation, financial resources and qualifications of staff will set certain limits to an organizing centre. In smaller organizing centres, the various functions may be undertaken by a few personnel, but it may be more efficient to delegate functions to external offices and professionals. Proficiency test items may be commercial or prepared locally and the organizing centre will have to consider these aspects.

4.1 Guiding principles

The written guiding principles (policies and regulations) of an organizing centre should include or concern, among other things, the:

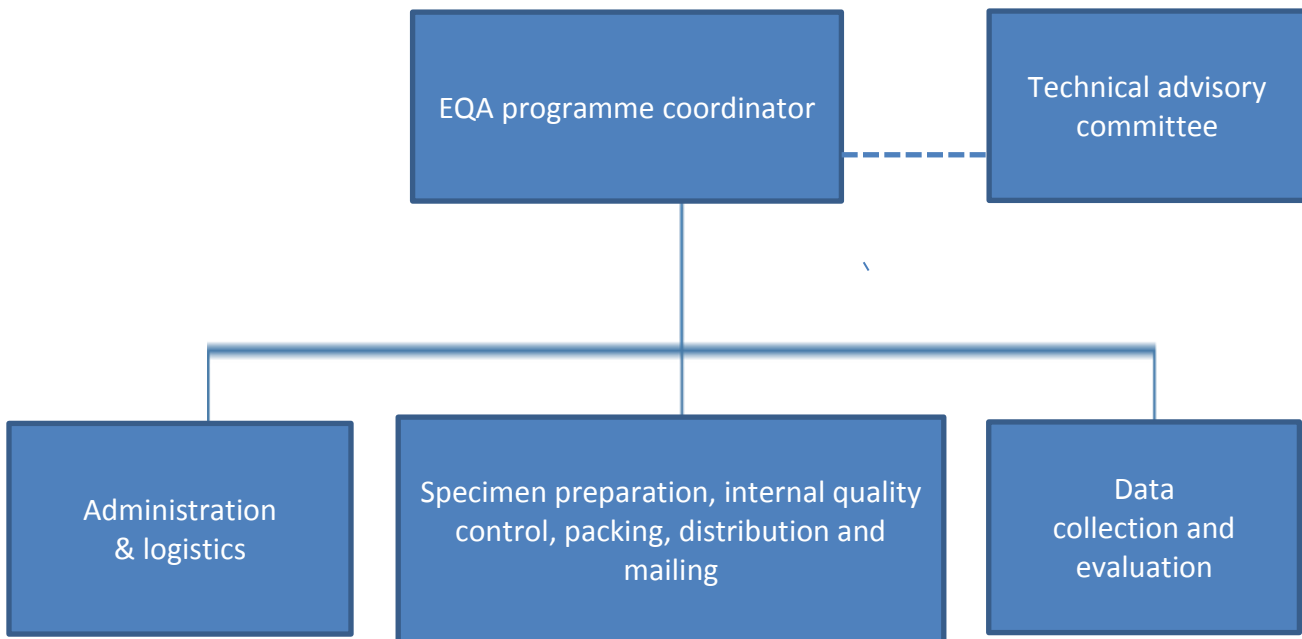
- provision of EQA services to interested parties (hospital laboratories, public and private laboratories, blood banks, other testing sites) through the conduct of regular PT rounds
- a flow diagram outlining the centre's structure;
- juridical identity, integrity and impartiality of the Centre;
- reinvestment of surplus funds in development of the programme where an organizer decides to charge fees from participants;
- confidentiality of providing information to participants related to their individual performances;
- use of results of PT rounds for educational and quality strengthening purposes, and improvement of laboratory services.

ISO/IEC 17043:2010 has a proposed general framework for the structure of an EQA organizing centre (12).

4.2 Structure

An EQA organizing centre should have a transparent structure; the example given in Figure 1 indicates the various sections and responsibilities.

Figure 1 Structure of an EQA programme organizing centre



4.3 Personnel

An EQA organizing centre requires managerial and technical experts to run the programme; the number and qualifications of staff will depend on the size of the centre, the planned PT round(s) and the activities. The following skills and functions are required:

- Programme director/coordinator
- Technical advisory committee
- Administration and logistics officer

- In-house quality officer
- Professional(s) in specific laboratory discipline(s) to prepare proficiency test items
- A statistician and computer expert
- Operational staff.

All staff should have a detailed job description and, where necessary, a deputy should be appointed (can be a member of staff with another primary function) to take over in case of absence. Not all positions need to be full or part-time; some can be shared or made available on a consultative basis e.g. statistician or computer expert. The continuing professional development of personnel should be documented, and experts designated as advisers for each laboratory discipline; the technical advisory committee is constituted to provide technical support. The organizer maintains records of all staff, as indicated in ISO/EIC 17043:2010 (12). Many countries will not have locally available personnel already trained and competent to coordinate, manage, or prepare proficiency test items, but opportunities for acquiring these skills can be gained through developing a relationship with an existing programme.

4.4 Room facilities

The following facilities are required:

- laboratory facilities to prepare and validate proficiency test items
- secretarial office (with telephone, fax, e-mail and copier facilities)
- space for packaging proficiency test items
- stock rooms with appropriate refrigeration for storage of proficiency test items
- secured area for data collection and evaluation
- archiving facilities.

All rooms should be restricted to authorized persons only, to avoid breaks in the confidentiality chain. Some proficiency test items, e.g. whole blood, may require cold chain management including an efficient transportation network. Proficiency test items containing potentially infectious materials, e.g. for microbiology or virology programmes, may require specialist courier services with required competencies; in some developing countries local courier services with appropriate competencies may prove to be cheaper and more efficient, therefore possible liaison with post offices, air services, etc. should be considered.

4.5 Equipment

Equipment is required for preparing, validating and storing proficiency test items. Specific equipment, consumables and durables may be required, depending on the type of analytes used, and the assays used to test these. All laboratory equipment should be calibrated according to the manufacturer's instructions, and preventive and corrective maintenance schedules established.

5 Planning and organization of proficiency testing rounds

For maximum efficiency, an EQA centre should strive to organize programmes for one or more disciplines, e.g. immunology, clinical chemistry, haematology, microbiology, virology. Another model is to develop nationally or regionally distributed centres, where one site has microbiology expertise and another chemistry expertise etc., which are linked through a common coordinating office. The preparation, distribution and testing of one or more proficiency test items, and analysis of the results of a test or groups of tests from a number of laboratories within a defined period of time, constitutes a PT round. The most successful programmes are those in which participants can expect regular despatch of specimens and fast return of results. Thus the coordination office should be organized for preparing or acquiring proficiency test items, arranging their packaging and distribution, receiving and analysing results, and returning results to participants in a timely way.

5.1 Planning for establishment of an EQA programme

Formation of a technical advisory committee is the first step in planning an EQA programme. This committee should comprise a representative group of stakeholders, depending on the scope of the EQA programme, including laboratory staff (managers, scientists, technicians), those who request testing (clinicians), implementing partners, other partners who may contribute to the testing, and a statistician. The planning should be undertaken in consultation with health authorities and laboratory directors. The technical advisory committee will assist in:

- planning the requirements of the EQA programme, including the: analytes to be used, frequency of distribution of proficiency test items, concentrations of the analytes, number of testing sites, scale-up plans, statistical design, and analysis;
- preparing instructions for participants, providing comments on technical matters raised by participants, and providing advice in evaluating the performance of participants;
- identifying and resolving difficulties encountered during the running of the programme.

The number of PT rounds per year depends on many factors. Generally, a higher number of PT rounds will provide more data points and enable better understanding of the performance of a testing site; a programme should strive to issue 2–12 rounds of PT a year (12). More than one analyte may be included in a PT round, depending on the availability of proficiency test items, the programme requirements and country situation. **Initially, PT rounds should be implemented with a small number of more frequently tested or higher priority analytes.** The number of analytes and their complexity may be progressively increased in subsequent PT rounds using a risk-based approach.

5.2 Planning for implementation of an EQA programme

In planning the schedule for PT rounds, sufficient time should initially be allowed to ensure that all components are in place, including for the preparation and shipment of proficiency test items. Methods for the preparation of proficiency test items should be established at least 6–18 months in advance of the first PT round, so that PT items can be prepared at the most optimal times and conditions. Once a programme is established, less time is required to plan the next distribution of PT test items or add a new analyte. Consideration should be given to factors such

as public holidays, vacation patterns and other factors that may influence the distribution, e.g. weather, postal peaks, transport systems.

After setting the timetable, a detailed plan of action should be prepared for each PT round. The quicker an EQA report is sent back to a participant, the more likely the PT report outcomes will be taken into consideration. A preliminary report, giving expected results without individual analysis, could be sent a day after the closing date to allow any necessary immediate action to be taken. A final, more detailed report, with analysis of a participant's results, may be sent later; however, production and return of each participant's report should be as short as practicable, ensuring that participants will not lose interest and alerting those laboratories showing unsatisfactory performance as soon as possible so that they can take corrective action. When all PT schedules are completed, they should be compiled into a master schedule. Adjustments to ensure that adequate staff and facilities are available may be necessary to cover the workload at different times. Table 1 gives an example of a time schedule.

Table 1 An example of approximate timelines for planning and implementation of a PT round

Stage	Check date*	Action
Planning	-18 months	Identification of laboratory disciplines and sub-disciplines
	-16 months	Decision on frequency of PT rounds in a particular discipline
	-14 months	Determination of dates of mailing of proficiency test items
	-12 months	Preparation or ordering of proficiency test items (for non-perishable stable items)
	-12 months	Ordering of packing materials
	-3 months	Confirmation of delivery of proficiency test items and packing materials
Implementation	-1 month	Control of the availability of proficiency test items Safety checks, adjustment of analytes
	-1 month	General review of preparatory work for the PT round Informing of participants about despatch and closing dates by post or email
	-3 weeks	Validation of target values and viability of microbiological specimens; preparation of questionnaire and marking key
	-2 weeks	Printing of addresses, forms, etc. In-house processing; aliquot dispensing and aliquot testing
	-1 week	Packaging of proficiency test items
	-1 day	Dispensing and packing of perishable EQA materials e.g. fixed blood specimens In-house validation of panels before despatch
	Day 0	Delivery for mailing
	+1 week	In-house validation of despatched proficiency test items
	+2 weeks	Closing date – this will vary according to the type of test (some are performed daily, some weekly, etc.), and to

		specimen transit and data return times
		Data entry completion (for manual data entry, two entries are advised)
		Data analysis and validation
	+3 weeks	Printing of PT round reports and generation of corrective action templates
	>3 weeks	Assessment of remarks, questions and comments
Update of the programme's historical records		
		Storage of leftover proficiency test items, where stability of analyte permits

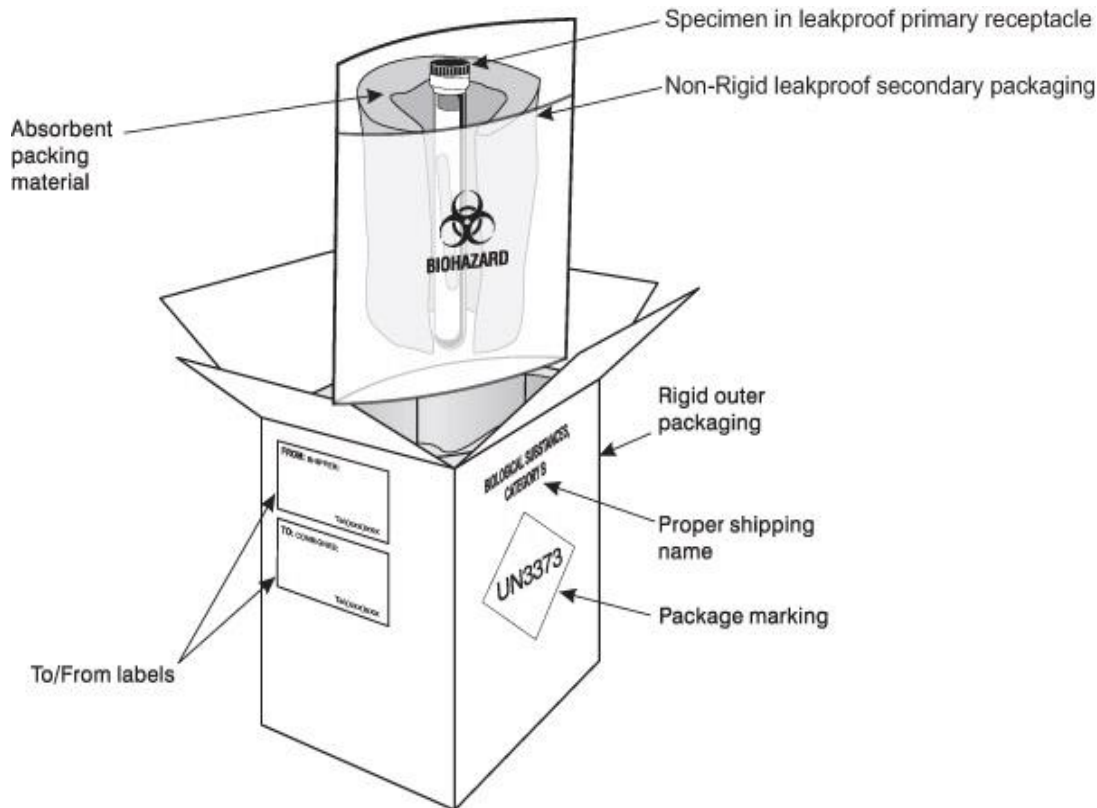
*The check dates will differ greatly depending on the type of EQA programme, number of participants, volume of data, analysis design, and availability of proficiency test items.

Reconstitution of lyophilized or dried proficiency test items: Specific instructions must be given for the whole process of reconstitution of lyophilized or dried proficiency test items. This includes the type and exact volume of liquid required for reconstitution, the time for warming up and equilibrating the vial, the temperature of the liquid to be added, and the time required for the lyophilized or dried PT material to dissolve in the liquid.

5.3 Packaging

The transport of infectious substances is regulated under the United Nations regulations for the transport of dangerous goods; infectious substances are Class/Division 6.2 (18). The type of packaging required for the transport of an infectious substance will depend on its classification under the regulations, on whether the specimen is classified as Category A, Category B or Exempt. Proficiency test items should be classified in order to determine the appropriate type of packing and transport. Category A infectious substances UN 2814 or UN 2900 packaging should follow packing instructions P620. Biological substances, Category B, UN 3373, packaging should follow packing instructions P650. Both Category A and Category B items require a triple package but each has different safety testing requirements. Exempt specimen packaging also requires triple packaging but with less stringent safety testing. An example of the triple package is illustrated in Fig. 2. The type of packaging, classification of the material, marking and labelling, and documentation of infectious substances must follow national and/or international packaging and shipping regulations (19). The type, number and volume of specimens to be distributed will influence the size of the package. Regulations and requirements for non-infectious materials are less stringent: the outer container should be labelled appropriately e.g. “perishable biological material”, and a specified set of documents should accompany the shipment, as per national recommendations and international regulations (18,19).

Figure 2 Triple package for transporting biological substances, category B



NOTES:

- Materials should be packaged in a fibreboard box or plastic box to prevent breakage during transportation.
- Glass or plastic vials containing lyophilized material should be packaged with cushioning material to prevent breakage during transport, and with enough absorbent material to prevent leakage during transport.
- Vials containing liquid material should be packaged in a leak-proof container and with sufficient absorbent packaging to absorb the entire volume of liquid.
- Primary and secondary containers should be leak-proof.
- Stabilized blood preparations should be transported under recommended conditions (e.g. frozen for sera, or in a cool box for fixed blood for flow cytometry).

5.4 Shipment

The international shipment or in-country transportation of proficiency test items must comply with the national and/or international regulations (18,19). A courier agent with the required competencies for transporting the proficiency test items should be engaged to work with the EQA organizer to determine optimal arrangements for despatch and timely delivery of the items, including notifying participants when the proficiency test items are shipped. It should be

ensured that local courier/postal services are aware of any recommendations for the safety of the courier and their environment, and, if applicable, of universal precautions for the use of dry ice and the transfer of dangerous goods and/or infectious substances (18). To control for the effect of transportation on the quality of proficiency test items, mimicking the transport conditions and examining possible effects on the specimens should be undertaken and any effects recorded.

5.5 EQA documents

The following EQA documents are essential; they should be developed and maintained:

- **Packing list:** A packing list is essential if more than one PT round is included in the distribution and not all participants have the same requirements.
- **Address labels:** Address labels should be printed from an up-to-date record of participants' contact names and addresses.
- **Instruction sheet:** Participants should receive adequate instruction regarding: proficiency test items characteristics and handling requirements, how to record results, where and how to return results, and the PT round closing date. Relevant notes and instructions should also be given about infectious agents for which proficiency test items have been screened, other hazard warnings, storage requirements, reconstitution method, recommended method for disposal, etc.
- **Protocol form:** Participants should be provided with a sheet to record results. The forms should be as simple as possible in order to minimize confusion in entering and reading the data and to facilitate data entry. Control data such as participant reference number, PT round number, proficiency test item(s) number(s), test method, etc. may be pre-printed if facilities permit, or can be completed by the participant. Pre-printing information will reduce transcription and omission errors. The participant's unique identifier should be placed on all protocol forms; this should be confidential to the participant and the EQA centre and not divulged to any third party. The units in which results are reported should be clearly shown. The protocol sheet may also include a list of codes (e.g. for methods and/or reagents); participants should enter the appropriate code for the method and/or reagent, and results of the measurement, into the protocol form, and should have space to note problems or other comments. The mechanisms by which results may be returned to the organizing centre (e.g. post, fax, email, website) should also be shown in the protocol form.
- **Return log:** It is essential to record the date of reception of each report at the EQA organizing centre, and if appropriate, the method of return, e.g. mail, mobile phone message, post, fax.
- **PT round record:** For retrospective control, the centre should have a standard hard copy or electronic record form, which should be completed for each despatch. Any problems encountered should be noted on the PT round record, which should be signed and dated by the member(s) of staff completing each item.

5.6 Data entry and evaluation

To ensure that participants receive their reports promptly, the methodology for data entry and analysis should be established and evaluated by the EQA organizer. Data can be entered

electronically into basic software, or into software developed for the type of analysis, in the following ways:

- manually, and performed independently by two persons, with decisions being taken on discrepancies;
- by scanning, which may be useful for check box protocol forms, as hand written figures can be problematic;
- electronically, through transfer of data entered on-line by the participant.

If data are being handled in any manner that requires transcription from one form to another, or from forms to a database, a procedure to minimize the risk of errors being introduced during the transcription process needs to be in place.

Approaches for evaluating PT round data are discussed in more detail in Section 9.

5.7 Proficiency testing round reports

The format of the PT round reports will depend on the method of analysis and computing facilities available at the EQA organizing centre. The PT round to which a report refers should be clearly shown; the report should indicate performance for each analyte being studied, and the deviation of the laboratory results from the overall consensus value and from the appropriate method (instrument) consensus mean. Ideally, individual reports for each participant should show their performance in relation to the total group of participants (without indicating any other laboratory information) along with an assessment of their performance. If multiple page reports are produced, each page should be identified as belonging to the report. Small programmes with limited resources can send the same summary report to all participants with individual notes attached to unsatisfactory performers. If there are laboratories that demonstrate unacceptable performance, the report must recommend that these are investigated and appropriate corrective action be taken or advice sought. Long-term performance for each analyte should be provided when appropriate. Reports including the results of the evaluation should be sent by mail, fax, or e-mail or through a dedicated website to which the participant has secure, password-protected access. The confidentiality of information should be assured. It is advisable to control the quality of transmission when information is sent by fax.

5.8 Records

It is essential to have a well-organized record system including for PT round records, data analysis and participant performance. It will be necessary at intervals to select information to be included in reports such as the annual report to participants, reports to expert committees and steering committees overseeing performance, etc. Additionally, data regarding the characterization and testing of the EQA materials should be well maintained; this can help in case of major discrepancy between the participating laboratories and testing sites and the organizing centre.

5.9 Archives

All information on PT rounds, participants and all quality documents should be archived either in hard copy or electronically as specified by national regulations. EQA participants should also maintain records of their EQA performance.

5.10 Documents for external use

Each participant should have access to the information and instructions on the services provided by the EQA organizer. The information should be written in a comprehensive way and in the local language.

5.11 Scientific materials

The organizing centre should prepare and distribute informative materials that address the key issues raised in the PT rounds. The materials might include a précis of current understanding and the relevance of the results to clinical care. If information is extracted from existing manuscripts, the centre must account for references and copyrights.

6 Cost estimation

6.1 Financial resources

The organization of an EQA programme requires specific investment and financial obligations to ensure a sustainable service. The resources may be provided from:

- an allocated budget from government authorities
- twinning with other EQA organizers
- fees from participants
- commercial suppliers
- academic centres
- long-term benefactors including trusts
- support from development partners as a short-term measure.

When support is given from commercial suppliers of reagents or instruments, a written agreement should be signed to ensure an independent management of the EQA programme with no obligations of the organizing centre vis-a-vis the supplier.

6.2 Costs of operation

Calculation of expenses should include, but not be limited to, the following items:

- Rent of premises.
- General costs (water, communication, electricity, insurance).
- Salaries for personnel or man-hours required.
- Information technology infrastructure and maintenance.
- Costs for maintenance and repair of laboratory equipment.
- Proficiency test items, which can be purchased or produced locally
 - costs of the proficiency test items (serum, plasma, urine, whole blood, etc.).

- investments for local production
- evaluation of proficiency test items (i.e. test kits, reagents, calibrators, consumables, etc.)
- assessment of stability and homogeneity of the proficiency test items.
- Packaging
 - packaging of proficiency test items
 - envelopes and labels for mailing.
- Administration of PT rounds
 - registration of participants
 - invoicing of participants (if applicable).
- Printing of forms, reports and catalogue.
- Mailing and/or courier costs
 - of proficiency test items
 - of reports.
- Evaluation of PT rounds
 - reimbursement of costs for experts (as appropriate)
 - costs for use of informatics (as appropriate)
 - costs for organizing meetings of experts and workshops for participants (once or twice a year).
- Training/corrective actions for participants.
- Miscellaneous costs
 - programme development costs
 - programme financial management
 - programme quality management system
 - staff training and development.

6.3 Subscription fees

Each EQA programme should be financially secure so that it can function in a timely manner and generate extra funds to support innovative development. Subscription fees from participants may be the only source of income. In determining the cost of the subscription fee, account should be taken of the number of participants and the number of PT rounds that are envisaged – the costs of implementation may vary for each PT round because the number of participants and number of PT rounds may vary each year, and there may be unforeseen challenges. Thus it is critical that the budgeting process is risk-based, and also takes into consideration the costs of quality control, research/development, and equipment care.

7 Proficiency test items

The materials considered for EQA should be as similar as possible to patient specimens. The starting materials may be of animal or human origin, microbial or artificial. They may include fresh or stabilized blood preparations; dried, liquid or lyophilized serum specimens; fresh or fixed, unstained or stained smears and films, or sealed wet preparations. Blood from a variety of animals can provide material for EQA purposes. Animal serum is widely used in clinical

chemistry and for special investigations in haematology. The materials should have no, or only minor, matrix effects with the analytical methods used. The compatibility of new equipment with EQA materials should be investigated by both parties.

7.1 Requirements of proficiency test items

The programme must ensure that the items are:

- as similar as possible to patient specimens whenever possible;
- homogenous, as indicated by homogeneity testing;
- stable at least for the PT round turnaround time (including proficiency test items transit conditions);
- safe and conforming to all relevant national safety standards, provisions and legislation;
- negative for infectious agents (e.g. hepatitis B and C, HIV 1, and HIV 2 in the case of human blood products) unless specifically required for the PT round;
- where applicable, ready for use with a pierceable septum if possible, especially when intended for use in an automated analyser, in order to minimize pre-examination errors and avoid opening of vials;
- sterile, except for specific instances such as microbiological PT rounds; this should be confirmed by testing random vials after dispensing – the number of vials tested will depend on the size of the batch produced and the conditions under which they are produced; and have
- acceptable matrix, homogeneity and stability properties.

These requirements are discussed in more details in Section 9 of this manual.

7.2 Collection and testing

It is convenient to use blood or blood components that have been collected by a blood bank where testing for HIV, hepatitis and other blood transmissible infections has already been performed. However, in some instances, the blood transfusion services regulations may require consent from the donor for the blood to be used for non-clinical purposes. If blood from a blood bank is not available, national regulations must be followed when obtaining other specimens. In some countries patient samples can be used for quality assurance purposes if they have been properly and completely de-identified.

Blood for transfusion purposes is collected into citrate phosphate dextrose anticoagulants, e.g. citrate phosphate dextrose adenine (CPDA), and should be as fresh as possible for EQA purposes. When volumes exceeding one unit are required, units of the same ABO group may be pooled. Whole blood should be stored at 2°–8°C.

To obtain serum, blood may be drawn into sterile containers with no anticoagulant, allowed to clot and the serum separated under aseptic conditions. In some cases, bulk volumes of serum may not be available for particular analyte testing (e.g. syphilis serology testing). Alternatively, plasma specimens can be converted to serum by de-fibrination. Coagulation factors present in plasma can be activated to form fibrin with the addition of calcium chloride and thrombin (20).

When blood specimens from different donors are pooled, each individual donation should test negative for hepatitis B and C and for HIV 1 and 2 as applicable, and should be tested before pooling. Smaller aliquots of serum, from up to 10 untested donors, may be pooled and the pool tested. Such pooled material is likely to require sterilization by micro-filtration. If for specific reasons material cannot be tested, it should carry a hazard label stating that it has not been tested for potential pathogens. This should also be stated in the participants' instruction sheets. Serum should be stored at -30°C or -40°C (or exceptionally -70°C); repeated freezing and thawing of serum should be avoided. Whenever pooled material is used, records should be kept showing the traceability of each component of the pool.

Material derived from animal sources should ideally come from a closed herd, certified free from diseases transferable to man and domestic animals. National regulations on the handling of animal materials should be considered. Similarly, when importing or exporting proficiency test items, international regulations should be respected. Biochemistry EQA materials can also be prepared by spiking specimens with specific analytes. Likewise, microbiological EQA materials may include pathogens or commensal bacteria or fungi. Regulations for the acquisition of all raw materials for EQA may vary from country to country, and relevant national guidelines should be consulted. Table 2 summarizes the advantages and disadvantages of using EQA material from various sources.

Table 2 Comparison of EQA materials from available sources

Source	Advantage	Disadvantage
Human	Similar to patient specimens	Poses risk of infection Sometimes difficult to obtain sufficient volumes May have greater turbidity Requires stringent ethical considerations
Animals	Limited risk of infection Available in sufficient volumes	Limited use because of possible unacceptable differences from human material Religious and cultural considerations may be necessary
Artificial (simulated)	Easy to prepare No known risk	Least similar to patient material Most likely to be associated with a matrix effect
Microbial	Ability to select specific pathogen or contaminant	Need to be formulated into clinically relevant specimens Increased risk of infection

7.3 Preparation of proficiency test items

The preparation of proficiency test items should be carried out in a dedicated area of the laboratory, away from the routine processing of specimens, using aseptic techniques and sterile glassware to minimize the risk of contamination. Agents such as antibiotics and antifungals or preservative (sodium azide) can be used in some specimens to reduce contamination. The proficiency test items should be stable for the whole period, from the collection date to the closing date of the PT round.

The dispensing volume of the proficiency test items depends on the number of tests and the volume required for measuring each individual analyte under routine conditions. The ratio of blood to vial size should be considered. Nearly full vials cannot be mixed properly while too small a volume may dry on the vial walls or deteriorate more rapidly. When additives are required, their amount in the proficiency test items should be kept as low as possible. The chemical name of the additive and its concentration should be mentioned in the instructions to participants or in associated safety information. Spiking materials should be added according to the analytical needs.

Microbiology proficiency test items should be tested seven days prior to the send-out date and be certified free from extraneous or unplanned bacterial or fungal contamination. If bacterial or fungal contamination is confirmed in one specimen, a second specimen should immediately be tested and if the results are positive the specimens should not be issued. Should any participant report bacterial or fungal contamination mid-PT round then a specimen from the retained stock should immediately be re-tested and, if confirmed, the PT round must be cancelled. The participant should return the suspect specimen to the EQA organizing centre for cross checking as contamination may have occurred in the participants' laboratory.

Vials for storing proficiency test items should be robust and leak-proof, and of a size that can be placed in a rack. They should be made of a material that does not leach ions (e.g. sodium, potassium, magnesium, iron) or otherwise interact, usually plastic or glass, and sealed tightly with a stopper that can be removed without tools. Each proficiency test item should be labelled to indicate its identity, including expiry date if applicable. In addition, the following information should be included in an accompanying sheet:

- type of proficiency test items (e.g. serum, stabilized whole blood)
- statement that the specimens are for proficiency testing use only
- instructions for appropriate storage (e.g. temperature: 2°–8°C)
- precise and clear instructions for reconstitution (if required).

This information may be included in the PT round instruction sheet or in an accompanying safety sheet.

Manufacturers providing proficiency test items to an EQA coordinating centre should also provide detailed information on the source of the material, additives, stabilizers and spiking material(s), and documentation of test results for infective agents, under a confidentiality agreement concluded between the two partners. The proficiency test items prepared locally by the EQA coordinating centre should ideally be verified by a reference or expert laboratory to ensure that the end product is consistent with the planned product and to ensure uniformity and adherence to standard procedures of preparation. Verification may be done at the beginning of the local programme and if possible at periods of 2 to 3 years, or more frequently if necessary.

7.3.1 Stability during transportation

Even if properly packaged, proficiency test items may be damaged by extremes of temperature, both high and low, or by prolonged transportation, or unduly rough handling. While all variations cannot be controlled, the EQA provider who sends out proficiency test items to participants should validate and demonstrate the stability of the items during transportation using the accepted guidance (12) during the planning stages of an EQA programme (see Section 9 for more details). At this time, a small number of participating laboratories, preferably those to whom delivery is expected to take longest, should receive two or more panels and be requested to return one panel unopened to the organizing centre using specified transport conditions depending on the type of specimen. The organizing centre will then test the panel again to ensure that the specimens have remained stable during the stringent conditions of twice the normal period of transport. If items are found to deteriorate during transportation then alternative distribution channels or type of PT round material might be considered. The organizing centre should, as a matter of routine, ship a set of PT items to itself for testing on the closing day of the survey and analyse these results alongside those of participants.

8 Broad considerations on the preparation of specific proficiency test items

Local preparation of proficiency test items is an important risk-based process that should only be carried out after considerable discussion, bearing in mind the efficiencies required and complexities involved. In this section only general principles will be discussed; the reader is advised to refer to specific guidance.

8.1 Haematology

Whole blood suitable for EQA of the blood count by automated cell counters is a challenge due to the rapid deterioration of fresh, anti-coagulated blood in vitro. Blood collected in ethylene diamine tetra acetic acid (EDTA), commonly used in patient specimens, starts to deteriorate six hours post venesection. However, blood collected in acid citrate dextrose (ACD) or citrate phosphate dextrose (CPD) anticoagulants has a shelf life of approximately three weeks (21) for the red cell parameters, and can be further stabilized by the addition of fixatives (22). Blood collected into ACD or CPD for white blood cell and platelet counts should be collected without leukodepletion and is unsuitable for use without stabilization unless distributed and analysed within 24 to 48 hours of collection. Commercially prepared, fully stabilized blood count material, similar to that used for internal quality control, has a relatively long shelf life; however, it is costly to buy and import and is generally specific to a single instrument type.

It is possible to prepare a haemolysate to which fixed platelets and avian red blood cells (“pseudo-leukocytes”) are added to make a PT material for control of haemoglobin, white cell and platelet counts, although the usefulness of this material for automated cell counters is limited. Haemolysate, to which a value for haemoglobin concentration has been assigned by a central laboratory, may have a use for the calibration and control of instruments used for haemoglobin measurement in peripheral laboratories.

Donor blood specimens usually have blood count values in the normal, healthy reference range. The values for haemoglobin concentration, red cell count and haematocrit can be manipulated during preparation by the addition or removal of plasma. If two pools of PT round material are being produced from ABO group compatible blood, plasma may be removed from one and added to the second pool to raise and lower the values respectively. Variation in white blood and platelet counts is more difficult but may be achieved by mixing of leuko-depleted and non-leuko-depleted ABO group compatible whole blood. Alternatively, non-leuko-depleted blood may be centrifuged gently to produce a buffy coat and this may be used to raise the white cell and platelet counts in one specimen, lowering the counts in the other at the same time.

Stabilized whole blood specimens are not usually suitable for automated white blood cell differential counts, as the method of stabilization may interfere with how the white blood cells react with the automated analyser reagents. Material for EQA of automated leukocyte differential counts should therefore either be fresh blood taken into ACD or CPD that is distributed and analysed within 24–48 hours of collection, or commercially prepared material. Stabilized whole blood is suitable for automated reticulocyte counting by most automated cell counters but will only have values in the normal reference range. Proficiency test specimens with raised reticulocyte counts are available commercially, and commercial material may also be required for certain analyser models that cannot process stabilized material produced by the EQA programme. Alternatively, fresh blood (as above for automated differential counts) may be used.

The provision of stained or unstained peripheral blood films is relatively straightforward, provided the EQA programme has access to a good range of patients' specimens and subject to ethical considerations. The EQA programme should aim to collect a "library" of disease conditions, from which material for distribution can be selected. The disease conditions for which blood films are prepared should reflect those seen in normal health care practice in the region, including blood parasites, occasional normal films, and some that challenge the participant. Unstained slides are used to assess the staining capability of a laboratory; in this case the slides are returned to the EQA programme for assessment of staining quality. Stained slides are suitable for the assessment of morphology skills and manual white cell differential counting. Slides of supra-vitally stained blood for manual reticulocyte counting can also be supplied, although the stain quality of these may deteriorate under prolonged exposure to extreme heat and humidity.

Lyophilized plasma is supplied as an EQA specimen for coagulation testing. The EQA programme must decide on the range of tests to be provided and whether there are different levels of participation, e.g. a first level for basic coagulation tests and a second level for more advanced investigations. Heparinized EQA control plasma is supplied to mimic specimens from patients on heparin therapy. Instead of reconstituting lyophilized heparinized plasma with water, a vial with heparin solution can be sent together with a vial of lyophilized normal plasma for reconstitution. In this way, different international normalized ratio values (INR) can be obtained from normal plasma specimens.

Where fresh material is to be distributed, all blood components must be screened for the presence of blood-borne infectious agents before use or be taken from individuals screened as negative for these infections at the time of donation. All haematology materials prepared by the EQA programme must be assessed for homogeneity, stability and sterility as described in Section 9.

8.2 Bacteriology

The selection of proficiency test items for bacteriology depends upon factors such as the prevalence of diseases, the tests being done in routine services and the cost and ease of despatching the items. Different types of EQA materials should be prepared for different levels of the health system, with perhaps less detail in reporting at each level. For peripheral laboratories, PT rounds may include acid-fast bacilli and exudate containing Gram positive/negative cocci/bacilli etc. For intermediate and central laboratories, most PT rounds are limited to the isolation, identification and susceptibility testing of bacteria. Bacteriology EQA materials may include clinical specimens with specific pathogens, pure cultures or mixtures of several clinically relevant organisms. The selection of cultures for a PT round usually takes into account the following:

- bacteria with public health or infection control importance
- newly recognized or recently renamed bacteria
- familiar pathogens with special antimicrobial susceptibility profiles
- the sending out of specimens containing normal flora without a pathogen, which can provide a valuable check on laboratory techniques and possible contamination problems
- for some PT rounds, presenting a challenge using bacteria with atypical characteristics may be of educational value.

EQA materials intended to provide clinical bacteriology laboratories with basic or complex challenges can be made locally, reducing the costs and complications of international transportation of samples that are designated as “dangerous goods”. EQA materials intended to challenge the ability to receive, culture, further test, and then appropriately interpret the results can be made with basic tools and simple materials. EQA providers may want to create simulated wound swabs, throat swabs, sputum, enteric (faecal) specimens or specimens for blood culture. To simulate these specimens, providers may want to consider either sending a pure culture, with or without a pathogen, or sending a more complex specimen that contains some organisms generally referred to as “normal flora” that are typical for the site. For example, a simulated throat swab might contain a combination of viridans streptococci as a background of normal flora, combined with *Streptococcus pyogenes* (group A streptococci). This would challenge a laboratory’s ability to detect beta-haemolytic streptococci in a milieu of viridans streptococci. Alternative pathogens can be used to replace the group A streptococci including group C streptococci, or group G streptococci. Alternate non-pathogens can be inserted such as *Staphylococcus aureus*, *Haemophilus influenzae* or *Haemophilus parainfluenzae*.

Alternatively, a simulated “carbuncle” specimen might contain a pure culture of *Staphylococcus aureus*, or a combination of *Staphylococcus aureus* plus *Staphylococcus epidermidis*, to

represent normal skin flora as one might see if the wound site was not cleaned prior to collecting the specimen. Other skin pathogens that might replace *Staphylococcus aureus* could include *Streptococcus pyogenes* (group A Streptococcus) as a simulated cellulitis or erysipelas, or *Pseudomonas aerogenes* as a simulated swimmer's ear cellulitis. The specimen can be made more complex by using methicillin resistant *Staphylococcus aureus* (MRSA). Simulated sputum specimens can be based on a substrate of egg white, or Mueller Hinton medium. Pneumonia pathogens may include *Klebsiella pneumoniae*, or *Serratia* species, or *Pseudomonas aeruginosa*. More fragile pneumonia-associated bacteria, in particular *Streptococcus pneumoniae*, may require supplementation with special agents that support growth, as described below.

As a cautionary note, trying to simulate a urethral swab that may mean working with fragile bacteria such as *Neisseria gonorrhoeae* may require a lot of research and development (R&D). This would be best avoided until the provider has developed strong R&D techniques. As an alternative to culture, one might consider creating a simulated slide for Gram stain, which contains either *N. gonorrhoeae*, or a more robust but less risky (in terms of risk to workers) bacterium that morphologically resembles it, such as *Moraxella catarrhalis*. Simulated enteric specimens can be based on a substrate of boiled brown lentil with a background of heavy concentration of *Escherichia coli* plus *Enterococcus faecalis*. Enteric pathogens can be added to the specimen in lower concentrations. Examples of enteric pathogens may include *Salmonella enteritidis*, *Salmonella typhi*, *Shigella* species, enterotoxigenic *Escherichia coli* or *Yersinia* species, or any enteric pathogen appropriate to the geographic region. Simulated blood specimens for testing in laboratory blood culture bottles can be provided with a single bacterial pathogen that has been attached to a latex bead along with 5 mL of citrated bovine whole blood. Other specimens such as simulated urine specimens or simulated cerebrospinal fluid can be created, however these may need supplemental amounts of bacteria stabilizer such as sodium formate and/or sodium sorbate. Creation of urine specimens requires additional work to ensure a concentrated stable specimen, and should thus be delayed until the R&D challenges have first been addressed. The concentration of bacteria in EQA specimens needs to be reasonably stable.

The growth rate of some robust bacteria, in particular Enterobacteriaceae and other Gram-negative bacteria, needs to be slowed using organic acids, as described above. On the other hand, some bacteria are more fragile, and require agents that support viability without necessarily supporting growth. Agents that can provide stability include bovine serum, egg albumin, or skim milk powder. Appropriate concentrations of these products will depend on the concentration required, storage temperature and required duration of viability, such that R&D may be required. Other EQA materials include simulated slides that resemble typical clinical specimens appropriate for Gram stain. To make these realistic, it is not only appropriate to include bacteria, but supplemental host materials such as serum, epithelial cells or neutrophils, can be added. Thus an R&D laboratory with capability is needed. EQA programmes are encouraged to start by creating specimens with an uncomplicated challenge. As the programme progresses in R&D ability, and as receiving laboratories become increasingly more comfortable with the PT process, specimens can become more complex by the introduction of less robust or

less relevant bacteria, more challenging susceptibility testing results, and/or more sophisticated interpretation of results.

8.3 Parasitology

Identification of parasites in tissues or excreta is fundamentally the same at all laboratory levels; it includes examination of stained blood films for malarial parasites, trypanosomes and microfilariae, and detection of parasitic cysts and ova in faeces and urine. Tests to identify parasite antigens in stool specimens, such as for *Cryptosporidium parvum* and *Giardia lamblia*, are also available. Tests to identify specific antibodies to parasitic infections in serum, such as *Toxoplasma gondii*, *Schistosoma* spp. and *Entamoeba* spp., are now commercially available, although most parasite serology (except for *Toxoplasma*) is performed in reference laboratories. EQA programmes are currently available for a number of parasite antibody tests. Antigen testing for malaria parasites using rapid diagnostic tests (RDT) is now commonly performed worldwide; stable antigen material for proficiency testing of malaria rapid diagnostic tests is under development.

Selection of parasites for EQA materials can be based on the disease epidemiology seen in a defined area. These include:

Blood parasites:

- Malaria: *Plasmodium falciparum*, *Plasmodium malariae*, *Plasmodium ovale*, *Plasmodium vivax*
- Microfilariae: *Wuchereria bancrofti*, *Loa loa*, *Brugia malayi*, *Mansonella* spp.
- Human African trypanosomes (*Trypanosoma brucei rhodesiense*, *Trypanosoma brucei gambiense*)
- South American trypanosomes (*Trypanosoma cruzi*)
- Relapsing fever: *Borrelia* spp.
- Babesiosis: *Babesia* spp.

Ova of intestinal helminths in stool:

- *Ascaris lumbricoides*
- Hookworm
- *Trichuris trichiura*
- *Taenia* spp.
- *Strongyloides stercoralis*
- *Diphyllobothrium latum*
- *Hymenolepis nana*.

Ova of flukes in stool:

- *Schistosoma mansoni*
- *Chlonorchis sinensis*
- *Opisthorchis* spp.
- *Paragonimus* spp.

- *Fasciolopsis* spp.

Ova of flukes in urine:

- *Schistosoma haematobium*.

Cysts/oocysts of protozoa in stool:

- *Entamoeba histolytica/Entamoeba dispar*
- *Entamoeba coli*
- *Giardia lamblia*
- *Balantidium coli*
- *Isospora belli*
- *Cryptosporidium parvum*
- *Endolimax nana*.

Parasite antigenic material for detection of specific antibodies (IgG, IgM) in blood/serum:

- *Schistosoma* spp.
- *Entamoeba histolytica*
- *Echinococcus granulosus*
- *Toxoplasma gondii*
- *Toxocara canis*
- *Strongyloides stercoralis*.

Specimens from patients are mainly used for preparing EQA materials for parasite identification, and are preferred over material from experimental animals. Parasites obtained from laboratory animals may not closely resemble materials obtained from human infections, but they are sufficiently similar to be useful. However, parasites grown in in-vitro culture cannot be used; for example, *Leishmania* organisms produce promastigote forms in vitro, but amastigote forms in human tissue. Antigens used for parasite serology testing are mainly obtained from animals. It is important to note that some parasite stages cannot be preserved, such as trophozoites of stool protozoa, for which only cysts or oocysts can be preserved in specimens prepared for EQA.

Three main procedures are used for the preparation of EQA material for parasites and parasitic infections:

- Preparation of blood films (stained or fixed, unstained)
- Preparation of preserved stool (or urine)
- Preparation of antigenic material for parasite serology (in reference laboratories).

Blood films are prepared from the blood of patients, but it is necessary to stain or fix the films before storage and despatch, as the morphology of parasites in unstained or unfixed films deteriorates. Unstained, fixed blood films are used also to monitor the quality of stains and buffers, and the staining procedures of participants.

Faecal and urine specimens from individuals can be pooled, emulsified in 10% formalin in saline and diluted if necessary before dispensing into vials. In sealed vials the specimens are stable for

several months. Careful specimen preparation is essential as small numbers of parasites may be unevenly distributed among the specimens. Negative specimens should be included in the EQA programme from time to time.

The EQA material should be accompanied by information on the history and physical examination of the patient, age, residence, history of travel, etc., which may be useful in the interpretation of the results. This information can be simulated to emphasize certain aspects of the infection and may not relate to the clinical information from the source patient. A simulated history is always required where material is obtained from animal sources.

Reporting on parasitology specimens is mainly qualitative, that is, presence or absence of the parasite, and identification of the species. Some reports may be semi-quantitative, such as reporting malaria parasite density in thick and thin blood films. Specimens for parasite serology are currently reported qualitatively as positive or negative for the specific antibody.

8.4 Clinical chemistry

Preparation of appropriate EQA material for clinical chemistry investigations is often considered to be simple and straightforward. However, choices need to be made, primarily with regard to specimen presentation, base material and additions. Lyophilized materials provide the greatest stability with storage time and tolerance of transport at elevated temperatures, but imported commercial materials are likely to be unaffordable for low and middle-income countries. The limited stability of liquid (frozen) specimens is likely to limit their applicability. Replacing 20%–25% of serum volume with ethylene glycol can stabilize even labile analytes such as glucose or enzymes for 9 days at 32°C or 12 months at –20°C. The relative considerations for serum specimens are summarized in Table 3 below; urine specimens are normally presented as liquid with preservative.

Table 3 Types of EQA serum specimens and their properties

Serum properties	Lyophilized	Liquid	
		Frozen*	Ethylene glycol
Convenience	Limited	Good	Excellent
Preparation	Reconstitution	Thawing, mixing	Mixing
Stability	Excellent	Limited	Good
Vial-to-vial variability	Dispense-related	Excellent (mixed)	Excellent (mixed)
Physical properties	More turbid	Unchanged	High viscosity
Clarity	Little changed	Good	Good
Cost	High	Low	Intermediate

* with or without preservatives (e.g. antibiotics, antifungals, azide)

Human serum is suitable for all analytes, and is essential for hormone, protein and other immunoassays. Serum stabilized with ethylene glycol is limited by its viscosity, which may lead to sampling problems with (semi-)automated analysers as it is incompatible with ion-selective

electrodes. For enzyme activity assays, the base material is less important than the additions. Isoenzymes with kinetic properties similar to those present in the circulation should be selected, and these may be of animal origin. Species is of little importance for urine specimens, but human urine is more readily available in large volumes.

The concentration of many biochemical analytes (electrolytes, metabolites, hormones, enzymes, drugs) may be increased by the addition of pure chemicals or other preparations. If this can be done quantitatively it provides an assessment of recovery of added analyte; however, it cannot yield concentrations lower than in the base serum. Freezing serum slowly in a narrow container and then allowing it to thaw slowly without disturbance results in layering, with concentrated serum at the bottom and diluted serum at the top. Careful sequential removal of fractions provides a range of differing concentrations. With care, these can be as wide as 50% to 150% of the starting concentration, and removing the uppermost 20%–25% permits replacement with the same volume of ethylene glycol. This yields specimens with higher and lower concentrations of serum proteins, and may also be used for other analytes. This is especially useful for lower concentrations, though these can also be attained by simple dilution; albumin solutions may be used for dilution to maintain total protein content and physical properties.

Specimens from subjects with specific clinical conditions are always desirable, subject to ethical considerations, and provide an opportunity for interpretative exercises. They may be essential for some analytes, including glycated haemoglobin and many immunology and allergy investigations.

8.5 Serology

Immunological/serological EQA programmes provide indirect evidence for the presence of an infectious agent and are intended for both quantitative and qualitative investigations. These tests aim at detecting an antigen or antibody. Serological systems that can be included in surveys include HIV (RDTs, enzyme immunoassay [EIAs]), syphilis: (venereal disease research laboratory test [VDRL]), *Treponema pallidum* assays (TP-EIAs, TP-RDTs, TPPA, TPHA), Anti-streptolysin O test (ASLO), Latex agglutination for *H influenza*, *N meningitidis* and *S pneumoniae* and lastly rapid confirmation/typing of the organism, e.g. *Salmonella*, *Shigella*, *Streptococcus* etc.

EQA materials for virological investigations (e.g. HIV, hepatitis B) may be infectious and should therefore be handled with extreme care. Certain EQA panels (e.g. HIV) may be distributed to assess the ability of laboratories/testing sites to detect variants of a virus at different concentrations. The panels should be well characterized by different methods before being distributed to participants. For safety reasons, HIV positive panel specimens should be heat inactivated (60 minutes at 56°C) and then filtered. HIV negative specimens should not be heat-inactivated as this may cause false positive reactions. There are currently several blood by-products that can be used as HIV EQA materials. They include serum or plasma, dried blood spots (DBS) and dried tube specimens (DTS). In some settings it is important to include HIV-1, and whenever possible HIV-2 and HIV-1/2 specimens, in EQA panels. Serum or plasma EQA specimens for HIV serology PT rounds can be obtained commercially or prepared locally. Bulk

volume HIV seropositive specimens from local blood banks that are routinely discarded are ideal for this purpose. They can be definitively characterized for HIV status, aliquoted and stored for future use.

8.5.1 Serum or plasma

Serology testing is undertaken using serum or plasma as a specimen type. An ideal source of large volumes of material for serology EQA programmes is plasma that is unsuitable for transfusion; this has the added advantage of having been screened for transfusion-transmitted infections including (as a minimum) HIV and hepatitis B and C. If discarded blood donations are to be used for other markers for which EQA is required, testing needs to be undertaken by the EQA organizer. When discarded donor plasma is not available or is unsuitable, specimens for serology EQA programmes need to be collected from clinical specimens, for which ethical approval needs to be obtained. Although plasma is an appropriate biological material for serology EQA, aspects such as limitations in volume or clot formation following refrigeration/freezing must be considered when it is used as an EQA material. While manipulations may be performed to increase the volume or quality of biological materials, it should be stressed that EQA materials should be as representative of “normal” specimens as possible, and that each manipulation will cause deviation from “normality”. A manipulation may also have an adverse effect on the results produced by different test methods. Thus, each method of manipulation needs to be extensively validated for its effect before applying it to EQA material production.

Heat inactivation of plasma and serum: The specimens used as proficiency test items may be inactivated to reduce the risk of contamination/infection of handlers. The heat inactivation process using a water bath (or incubator) at 56°C for 30 minutes is adequate to inactivate many blood pathogens except hepatitis B or C. However, EQA panels should still be handled as potentially infectious and all precautions required for handling potentially infectious blood specimens should be observed. NOTE: The temperature in the water bath should be monitored with a thermometer throughout the time required for this process.

Conversion of plasma to serum: Plasma can be converted to defibrinated plasma (in effect serum) by artificially initiating the clotting cascade through the addition of thrombin followed by removal of the resulting clot. Once the clot is removed, microfiltration of the serum should be considered to remove any bacterial contamination.

Filtration: Biological materials can be filtered using either vacuum or pressure filtration apparatus, in a stepwise reduction of pore size (pre-filter, 0.8 µm and 0.45 µm and a final 0.22 µm). Filtration will remove any micro-particulate matter, ranging from micro-clots to bacterial contamination. As a note of caution, the conversion of plasma to serum and the filtration process are cumbersome, labour intensive and potentially hazardous, and should be conducted in a biohazard cabinet by personnel with adequate personal protective equipment including gowns. Without conversion of plasma to serum, reasonably satisfactory PT materials can be produced by centrifugation of the material to remove particulate matter and clots. To maintain

sterility without filtration, PT specimens should be handled in a Class II biosafety cabinet, paying attention to aseptic techniques to the extent possible and using sterile equipment and containers.

Pooling: For large EQA programmes, pooling can be considered an option for increasing the volume of material available (where the volume requirement for the number of participants exceeds that normally found in a single blood donation). Pooling specimens involves mixing individual specimens together to create a mixed pooled specimen. However, in pooling specimens, a dilution effect may be created; in other words, dilution of the individual constituents of each specimen used in the pool. This risk may be reduced by selecting specimens of similar antibody and/or antigen profiles, as ascertained by testing, thus only pooling specimens with similar test result profiles or characteristics. Pooling has the benefit of extending the volume of material while seeking to maintain the reactivity of key constituents in the specimens contributing to the pool. Sometimes, pooled negative sera or plasma can show an increased tendency to false reactivity. Therefore, pooled materials, whether they are negative or positive, should be tested extensively before use in an EQA programme.

Dilution: Dilution of biological material to increase the volume may be possible if the material is diluted only to the extent required to create the additional volume. Dilution to mimic early infection antibody profiles or “weak” results creates a final material that does not have an antibody profile representative of that seen in undiluted specimens. In addition, the ability of some tests to detect diluted antibodies and/or antigens that would normally be detected when undiluted may be reduced. The EQA programme organizer must test diluted materials in a large number of assays that are likely to be used in the programme to ensure that they behave as expected. Dilution of quantifiable markers (e.g. hepatitis B surface antigen [HBsAg] or HIV-1 p24 antigen) with negative serum/plasma may be considered to produce serological EQA materials representative of early infection, where it is difficult to find naturally occurring biological material of sufficient volume. In the case of HBsAg, the diluent must be free of hepatitis B surface antibody as it is likely to complex with HBsAg and render it unavailable in the testing system.

8.5.2 Dried tube specimens

In developing countries, HIV testing using rapid diagnostic tests (RDTs) is commonly offered outside the traditional laboratory setting and by non-laboratory professionals through task shifting, resulting in an unprecedented increase in the number of testing sites. To adequately monitor the accuracy of the HIV results, it is critical that efforts are made to ensure participation of the HIV testing sites in the national EQA programme. However, EQA programmes using conventional serum or plasma specimens as described above may be costly and difficult to implement in these settings; this type of specimen often requires stringent conditions for storage and transportation. Thus, to address the implementation challenges, an approach using dried tube specimens (DTS) (figure 3) – innovative, simple and easy to use – was developed to monitor and improve the quality of HIV serological testing in low and middle

income countries (see more details below) (14). Once rehydrated, DTS can be tested by RDTs or EIA, and therefore can be used by all laboratories and sites performing HIV testing.

The DTS preparation method has several advantages. It is safer and less biohazardous than preparing liquid specimens, and the specimens are stable at 37°C, including during storage and transport, and can be transported at room temperature without the need for maintaining an expensive cold chain. Once received at the testing facility, the specimens can be stored at room temperature for up to 4 weeks without negatively affecting their integrity. This new approach uses a 10-fold smaller volume of specimens than most EQA programmes, which use specimen volumes of 0.2 mL to 0.5 mL.

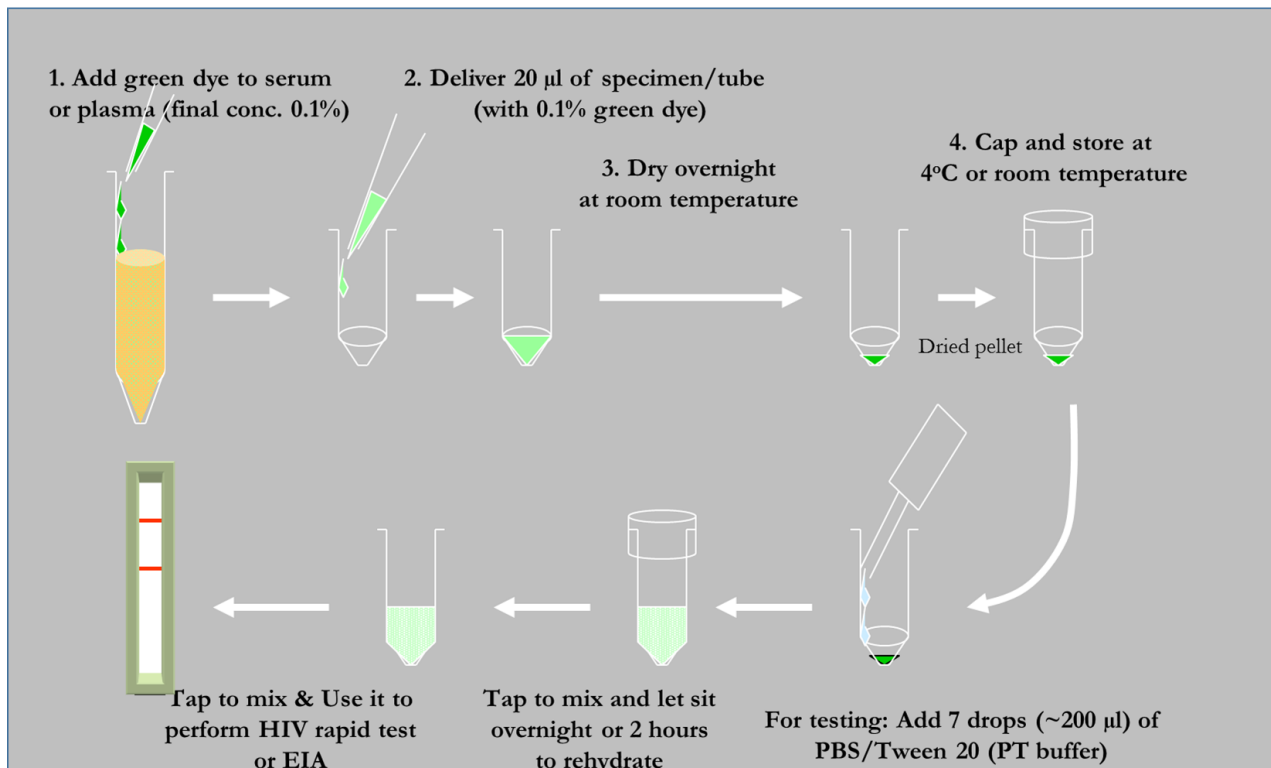
Preparation of dried tube specimens: The methodology for preparing DTS of HIV, syphilis and hepatitis using characterized serum or plasma specimens for control has been well established; it is summarized in Figure 4 (14). In short, dried tube specimens are prepared by transferring 20 µL of serum or plasma, premixed with 0.1% (vol/vol) green dye (i.e. food colour), into a 2-mL cryovial. The addition of 0.1% green dye does not affect the HIV test results but allows visualization of the coloured pellet at the bottom of the tube. The tubes are left open in a laminar flow hood at room temperature (18°–25°C) overnight to dry. The following day they are capped and stored at 4°C until rehydrated prior to testing. Before testing, the DTS are rehydrated with phosphate buffered saline (PBS) containing Tween 20 (PBS/Tween 20) buffer. PBS–Tween buffer is prepared, filtered through a 0.2-µm filter and aliquoted in 1.5 mL volumes to be used as rehydration buffer (also termed PT buffer). A day before testing, a DTS specimen is rehydrated by adding 200 µL PBS-Tween with a precision pipette, or 7 drops with a plastic disposable transfer pipette (catalogue No. Fisher Scientific, Waltham MA, Cat # 13-711-43).

NOTE With the above-mentioned pipette, seven drops amount to about 200 µL, but the number of drops will vary depending on the type of disposable transfer pipette used. It is recommended to validate the pipette by determining the number of drops that will yield approximately 200 µL. This results in a 1:10 dilution of the original specimen but is treated as undiluted for the purpose of further testing. The specimens are then mixed by gentle tapping, without vortexing, to mimic practical conditions encountered at the testing sites with limited or no laboratory equipment. The reconstituted specimens can be left overnight at room temperature to allow solubilization of dried serum/plasma into the PT buffer. Two hours later or the following day, the specimens are again mixed by gentle tapping and used to perform rapid tests or enzyme-linked immunosorbent assay (ELISA).

NOTE For laboratories and testing sites performing EIA/ELISA or Western blot, it is important to ensure sufficient EQA materials are provided to perform the approved testing algorithm.

Figure 3 Summary procedure for preparation and testing of DTS specimens

DTS Preparation and Testing



8.5.3 Characterizing EQA serology materials

The materials constituting proficiency test items in serology EQA must be homogenous and stable for the period over which the exercise is underway; and the true status of the materials must be known for the markers under assessment.

Characterization of EQA materials is usually achieved using a testing strategy and a defined algorithm. A testing strategy defines the range and types of assays used to determine the status of a specimen; an algorithm specifies the names (brand) of the assays and the order in which they are used. A testing strategy includes a first assay to identify negative specimens, and one or more supplemental assays that are used to confirm whether those specimens that are reactive in the first assay are positive or not. This confirmation of positive reactivity is performed because all serology assays give a small proportion of false positive results. Testing in a second assay assists with ensuring materials that are assigned a positive status are truly positive. The first assay in an algorithm must be of high sensitivity, especially if it is used to identify negative specimens, i.e. if a negative result from the first assay will be assumed to be correct and no other testing will be performed to “confirm” the material’s negative status. Therefore the EQA provider needs to be confident that the first assay used to characterize EQA materials is sensitive and available, and will not misclassify positive specimens that could be

detected by other assays used by laboratories participating in the EQA programme. The EQA provider needs to consider carefully the way materials are characterized. For example, if any participating laboratories or testing sites indicate that they use HIV combined detection assays, which detect both antibodies to HIV and HIV p24 antigen, the EQA materials must be characterized for both of these markers. Similarly, if participating laboratories use specific and/or non-specific treponemal tests for syphilis, the status of the EQA materials for both of these markers should be known.

8.6 Nucleic acid testing

Plasma, dried blood spot (DBS) specimens, and DTS are well-established specimen types used for nucleic acid testing (NAT) EQA programmes (23,24). Whole blood collected in EDTA collection tubes is acceptable, and can be spotted to create DBS or centrifuged to separate plasma. HIV positive specimens should be heat inactivated for 30 minutes at 56°C and then filtered prior to preparation of the panel. Laboratory isolates can be diluted using negative human whole blood or plasma when necessary; however, HIV positive (and negative) human specimens are preferred.

Panel target concentrations and viral subtypes/genotypes: When creating panels for molecular assays, it is important to include specimens across the linear range of the predominant assays being applied in the EQA programme; for example, one or more of each concentration: HIV RNA not detectable, HIV RNA detectable (low level) positive, and HIV RNA detectable (high level) specimens should be included in each panel. Panels ideally should contain 5 specimens, and either duplicate panels or an additional volume per tube should be included in each shipment so that laboratories have extra volume/aliquots of each specimen to perform an additional run in the event of instrument failure. If possible, viral subtypes or genotypes in the EQA panel should be representative of local subtype diversity. Including duplicate specimens in at least one panel per year, e.g. duplicate low-positive specimens in one panel, is recommended to evaluate reproducibility.

Plasma or DTS panels are important for HIV NAT assays that require liquid specimens (qualitative or quantitative). Plasma and DTS panels are available commercially. DTS is currently offered at lower costs or in some cases free of charge (24). Plasma and DTS panels are produced as previously described (24). Briefly, either high titre patient plasma is pooled and diluted or, alternatively, established HIV strains from standardized programmes are grown in the laboratory and then diluted in HIV negative plasma.

Other qualitative or quantitative assays require DBS specimen panels. For early infant HIV diagnosis (qualitative), DBS EQA panels are currently available at low cost or free of charge from a few global providers. DBS viral load panels (quantitative) are not yet widely commercially available. Production of DBS viral load panels using specimens has been described (25). DBS EQA panels are produced by spotting filter paper cards with pooled and pre-diluted HIV positive whole blood, or 8E5 cells with HIV viral isolate diluted into negative whole blood as described (24,25). The DBS cards should be allowed to air dry in a biosafety cabinet for at least four hours before packaging, and stored with humidity indicator cards and desiccant. DBS cards for NAT

can be shipped at ambient temperatures but should be stored at -20°C upon receipt if testing will not be performed immediately. DBS for molecular assays can be stored at -70°C for up to 2 years.

Considerations for shipping: Plasma requires a strict cold chain including dry ice shipments, which is often cost prohibitive. Both DBS and DTS can be shipped at room temperature.

8.7 CD4+ T lymphocyte testing specimens

In CD4+ T lymphocyte EQA PT rounds, both percentage and absolute number of CD4+ T lymphocytes per μL of whole blood are determined. More than one level of CD4+ T lymphocytes should be included with each trial and should typically encompass specimens with CD4+ T lymphocyte counts at appropriate clinical decision levels, i.e. a decreased CD4+ T lymphocyte count and a “normal” or increased CD4 count.

Fresh whole blood specimens collected in EDTA are the best EQA specimens to use if they can be tested within 8–24 hours of collection by laboratories; this is possible when the number of testing centres is small and they are close to the EQA provider. As this may be possible in only a few situations, CD4 EQA materials are usually fixed blood specimens and their production is complex. Typically specimens are made from artificially CD4+ T lymphocyte depleted, stabilized blood to avoid the risks associated with working with HIV+ specimens (26). If working with HIV+ specimens, the EQA organizer must note the International Air Transport Association (IATA) regulations that cover the transport of pathogenic specimens (19). Blood units to be used in the production of EQA samples should be screened for suitability taking into consideration the integrity of all relevant cell populations and the absolute CD4+ T lymphocyte count.

These materials are therefore best obtained from commercial sources or collaborations with existing international programmes that utilize such materials, either producing them in-house using patented or industrial secret approaches or using materials from commercial sources. However, the matrix of the material should be stabilized whole blood and be shown to be compatible with all flow cytometric platforms wherever possible.

When pooling multiple units of blood, the target CD4+ T lymphocyte count and blood group compatibility of the units must be taken into consideration. When using artificially CD4+ T lymphocyte depleted material, it is important to ensure that the CD8+ T lymphocyte count is not elevated beyond levels typically seen in a clinical scenario. This can be mitigated by using filtered blood units to increase the volume. Prior to aliquoting EQA samples, it is important to mix the EQA material adequately to achieve a homogeneous suspension. Samples should be tested for suitability prior to sending out to participants. Not all stabilized materials may be compatible with all CD4⁺ T lymphocyte enumeration platforms (including newer flow cytometers and point-of-care [POC] devices), and this should be taken into consideration. In such instances, alternative platform-specific material may be used and the particular user group monitored outside the main group as a separate PT round. Consideration should be given to the minimum number of participants required to generate valid statistics in such instances. The

final consensus EQA report should convey and score participants' results according to z-scores or 2 standard deviation (SD) limits and indicate if the results are acceptable or not, with a corrective action template supplied if any result is outside the 2SD limit.

8.8 Tissue-based pathology

Preparation of good quality tissue EQA specimens is fundamental to any EQA programme. Poor tissue preparation will make it difficult to assess any given biomarker using immunostain or molecular techniques. The selection of specimens is dependent on the test to be evaluated. Mostly likely, this will be formalin-fixed paraffin embedded (FFPE) tissue such as archived diagnostic material or banked material. Specimens of human origin should be used where possible; however it may be difficult to source some tumour types, as tumours can be heterogeneous (i.e. uneven in their tissue composition) and limited in amount. Selection of normal tissue for use in an EQA programme is a reasonable alternative in many, but not all, circumstances. Formalin fixed cell lines can also be used; however these will not necessarily reflect the clinical characteristics of tumour tissue composition.

As a general principle, EQA tissue specimens should be well characterized before use; this means validating the materials by both pathological review and molecular analysis (where applicable). When cutting slide sections from a single block, histology and target antigenicity should be reviewed at intervals throughout the process as differences in the specimen may occur in consecutive sections of the tissue block. In addition, for molecular pathology, consideration should be given to the percentage of tumour cells in the specimen, as sufficient tumour content is required for analytical sensitivity of many molecular tests and will vary depending on the test being used (27). Applying best practices for DNA extraction from FFPE material will help ensure good quality nucleic acids for use in molecular testing (27).

All tissues should be fixed in 10% neutral buffered formalin (pH 7.2–7.4) at a minimum ratio of 20:1 (fixative to tissue); a minimum fixation time of 18–24 hours is advised (28). The use of decalcified material in EQA programmes should be avoided unless assessment of a specific antibody/decalcified tissue combination is required; decalcification should only be performed on well-fixed tissue. Positively charged slides or treated slides (poly-L-lysine or saline) should be used to ensure maximum tissue adhesion. The 4 µm tissue sections are floated on clean (preferably distilled) water with no additives, and the slides are dried overnight at 37°C and labelled appropriately. In order to ensure continuous availability of EQA specimens, it is advisable to bank tissues. Data for banked FFPE tissues should include: date collected, tissue type, ischaemic time (time from collection to fixation), and total fixation time (time from initial exposure to fixative to tissue processing).

9 Requirements during evaluation of EQA results

The scientific validity of laboratory tests is limited by uncertainties at the biological/nosological level as well as by analytical errors. It is therefore important that a laboratory is able to identify

whether the difference in observations is indicative of biological variations or of technical differences in measurement. For practical reasons, one can distinguish two basic types of laboratory investigation: quantitative and qualitative. The results of qualitative tests are descriptive (e.g. blood group antigens and antibodies, blood cell morphology); the results of quantitative tests are numerical (e.g. haemoglobin concentration, serum enzyme activities, creatinine concentration). Some investigations combine descriptive with numerical information (e.g. differential white blood cell count, protein electrophoresis). Semi-quantitative investigations provide graded results that are less accurate and less precise than quantitative assays, but are still useful for clinical diagnosis and monitoring (e.g. urine dipsticks, malaria parasite counting).

The analysis of PT round results has three basic aims:

- To provide an overall summary (total number) of those in consensus and those out of consensus.
- To provide for each individual laboratory an analysis of its performance in the current and previous PT rounds based on the performance criteria established and defined for each analyte/programme.
- To identify possible causes of deviation of results when possible.

For the EQA provider, there are a number of considerations. For any given send-out, the EQA provider is expected to assure all laboratories that:

- the challenge they have received is consistent with the planned value (**controlled sample**);
- they have received the same challenge material (**homogeneity**);
- if there is a predetermined correct target value, laboratory results have provided evidence that the challenge is fair and reasonable to the extent that all laboratories can be expected to achieve a valid result (**equity**);
- where collective laboratory results form the basis of a consensus value, laboratory results will be compared and analysed in a fair and equitable manner (**comparability**);
- the probability or risk of specimen degradation has been studied and, to the extent possible, contained (**stability**).
- For review of proficiency testing of quantitative values, measurements in a laboratory should not be considered as perfect; some variation may be introduced into the testing process (**measurement uncertainty**).

Each of these considerations is discussed in more detail below and must be monitored in a manner that can be documented and shared with others as required.

9.1 Controlled sample

Each step in the production of proficiency test items must be monitored by quality control (process) activities that are appropriate to the type of specimen being produced. If the planned specimens are to be measured for a quantitative or semi-quantitative value (concentration, numerical count, weight), all reagent and material components must be measured to ensure

they are consistent with the amounts required by the product production plan. All equipment to be used in the production process (e.g. scales, centrifuges, incubators, refrigerators, freezers, spectrophotometers, pipettes, analysers, gas concentrations) must be known to be working properly; it must be monitored regularly using assured standardized materials to demonstrate and document that it is performing accurately. Equipment demonstrating inappropriate variation from expected values must not be used until the sources of error are found and corrected and the equipment is appropriately recalibrated.

9.2 Homogeneity

At the end of production, but prior to preparation for distribution, specimens should be tested to demonstrate they have a consistent value within the tolerance limits based on the measuring capability of the method or the anticipated evaluation. If specimens can be analysed in a manner that will not destroy them during testing, specimens can be randomly drawn from the total number of products produced. For absolute assurance that the produced lot is acceptable, all specimens could be tested. This however may represent a large burden of time or testing cost, so specimens should be selected from the beginning, middle, and end of production. The number of specimens to collect depends on how many specimens are produced. For lots of less than 100 specimens, a minimum of 10% of specimens should be chosen for testing; if fewer than 30 specimens are produced, a minimum of 3 specimens should be chosen.(17) The specimen results should be reviewed visually for reasonable agreement, and the standard deviation of the results calculated. Ideally this standard deviation will be less than 10% of the anticipated total evaluation limit, or less than 30% of the target for the participant standard deviation (17). If it is difficult to calculate the standard deviation, a rough estimate is the highest result minus the lowest result, divided by 4 (17). Any results that are vastly different from the majority are evidence of poor homogeneity, which can be verified by testing more specimens; alternatively, the production lot can be scrapped.

A variety of statistical techniques for homogeneity testing are available in the International Standard ISO 13528:2015 (17). These techniques cover different designs, the ideal design being to select 10 or more specimens randomly throughout the production run and to test each specimen in duplicate. Statistical spreadsheet software for analysis of variance is used to calculate the SD between specimens, which can be tested as discussed above (SDs <130% of anticipated [or actual] SD between participant results).

NOTE Many EQA providers do not produce their own specimens, but rather purchase the specimens from another supplier. Nonetheless, providers must ensure that the specimens meet the requirements for homogeneity; they can meet this obligation by:

- requiring the supplier to provide a certificate attesting and documenting that the purchased lot is tested and meets the requirements for homogeneity as given in ISO 13528:2015 (17);
- removing a small randomly selected number of specimens from the purchased lot and testing them for value, or sending them to an appropriate reference laboratory to have the tests performed.

9.3 Stability

Laboratories need to be assured that the specimens they are testing have a low probability of degradation between the time of production and the conduct of laboratory tests. By demonstrating specimen stability, variation that results from transport and storage is less likely to impact a laboratory's performance in the proficiency test. Transport can have a particularly challenging impact on specimens because of potential variability in the duration and conditions of transport. For example, specimens that are transported by road in certain geographies may be exposed to temperatures that exceed 40°C, while specimens that are transported by air may be exposed to temperatures below -40°C. It is important that the packaging reasonably protects the specimens from extreme conditions during transportation.

Specimen stability is a particular problem for biological specimens because of the presence of degrading enzymes such as proteases and DNAses. Specimen stability can be enhanced by the use of chemical stabilizers, or reduced temperature, or lyophilization; however, these methods may sometimes interfere with the actual testing process. Stability testing requires random specimens to be selected at various points in time after production to ensure that the results have not drifted due to specimen degradation. Often these occasions include (a) immediately post production, (b) at the time of specimen despatch, (c) at the time of the estimated or pre-determined date when laboratories are expected to perform specimen testing, and d) on the last date when laboratories can perform proficiency testing. The extent of testing will depend on the nature of the specimens and the provider's resources, but at a minimum the specimens need to be tested at points b) and d) mentioned above, to reflect any changes that can occur during the actual conduct of the proficiency testing.

Depending on the source of the materials, the results of homogeneity testing can serve as the initial stability testing point, and as few as 2 or 3 randomly selected specimens can serve as the stability test at the end of the test period. The criterion for acceptance of sufficient stability can be determined by the difference between the mean of the homogeneity test data (pre-shipment) and the mean of the results on the stability of specimens at the end of the testing period. The criterion recommended in ISO 13528:2015 (17) is that the difference be no larger than 10% of the error criterion, or 30% of the participant SD. If more testing points are used, then the test would be for a trend, or any evidence of loss of viability before the end of the study. In the case of qualitative results (presence or absence, or identification of a species), the criterion would be based on evidence of continued integrity.

NOTE Some specimens may be found to be stable for only a very short time (arbitrarily less than 7 days), so not allowing sufficient time for sending and testing the specimens; such specimens are not appropriate for sending out. Others that are stable for periods of (again arbitrarily) 14–21 days can be considered appropriate for sending out in regions where transport can be completed in 5–7 days. Specimens with very long stability are safe for sending out in all areas. During the phase of new specimens, it is appropriate to perform transport testing studies to determine if the samples are likely to be impacted by transport conditions. Once these have been studied, it is usually not necessary to repeat transport testing studies with each send-out because the prior gained information should suffice.

9.4 Equity

The ability of a laboratory to correctly interpret a proficiency testing sample may depend upon the complexity of the sample, the impacts of timing and transport and stability factors, and the skill set of the laboratory staff. If the challenge is such that a set of known reference expert laboratories or the collective group of all laboratories is unable to agree on the “correct” interpretation of the sample, the sample may be considered inappropriate for grading. In order for a sample to be considered acceptable for assessment and grading, some programmes require the “correct” response to be recorded by at least 80 per cent of the reference expert group and/or at least 50 per cent of the total group.

NOTE If a provider repeatedly sends out samples that do not meet acceptable equity levels, the point at which inappropriate selections are being made needs to be determined.

9.5 Comparability

At the outset, the EQA provider needs to determine the desired comparability of results for the proficiency test. The most important determination is whether the results should agree with a reference result (such as a “true” result or a result based on agreement of an expert group) or if the results are going to be compared with those of other participants. Reference results are generally more reliable for most analytes, but they can also be difficult to obtain, or expensive. Most proficiency testing is performed by comparing results among participants, often with participants using the same test method. Secondly, the EQA provider must determine the extent of comparability that is needed for a laboratory to be evaluated as having “acceptable” performance. Again, is the criterion to be based on a medical or technical expectation (e.g. 10% for glucose in whole blood), or will it be based on the extent of agreement with other participants? The design chosen by the EQA provider leads to the statistical method used to evaluate performance.

The simplest design is where a reference value is used as the target value and there is a pre-determined criterion for performance, such as within an analytical goal percentage from the expected value, or in qualitative results, having the correct result. The results are either correct or not (or within a predetermined range of the reference). It is possible also to have a target value chosen as the average of participant results, but have a predetermined criterion for performance based on analytical needs, regulatory specifications or clinical relevance. Similarly, the EQA provider may choose to use a target value determined by a reference lab (or a known source), but have the criterion for performance based on agreement among the participant group.

The most common design for proficiency testing is to determine the target value and performance limits entirely from participant results. There are very different approaches, numbers of necessary results, and statistical considerations for the different designs. Some of the possibilities are quite simple, requiring only one participant (having a reference value and a

pre-determined limit), and other designs will require 50 or even more participants (determining mean and standard deviation of participant results when different test methods do not agree). The full range of possibilities, statistical techniques, and statistical considerations, is given in ISO 13528:2015.(17) This manual discusses only the simplest techniques for calculating the robust mean and standard deviation of participant results, and the commonly used z-score performance statistic.

NOTE The need for a wide variety of statistical techniques is due to the wide variety of analytes, test methods and designs used for different proficiency tests. Most common summary statistics are based on an assumption of a normal distribution of results (a “bell shape” curve if graphed as a histogram). While this is the most common underlying distribution for laboratory test results, it may not always be the case. In every set of proficiency test results there are outliers, or other results far from the group; these values can severely skew (or bias) the summary statistics, and so must be controlled. These results can come from many sources – most commonly from laboratories that made a mistake, but can come also from poorly worded instructions, poor test methods and contaminated or unstable samples. In other cases the nature of a measurement produces highly skewed distributions of results, for example in organism counting tests; in these cases the data need to be transformed to get a symmetric distribution, or method groups need to be separated.

In almost all cases, the objective of statistical techniques is to produce an estimate of a population mean and SD. In different situations, this might be the mean and SD for all competent laboratories, for all laboratories in a defined area, or for all laboratories in the EQA programme that followed the instructions. The statistical techniques are the same in all cases but the designs and interpretations can vary. Outlier detection techniques are available in the statistical literature, but should be conducted only by individuals who understand statistical analysis, to avoid major errors that lead to incorrect performance evaluations. There are also several statistical techniques called “robust techniques” that are not severely affected by outliers, and can be applied in the common proficiency test situation with a mix of results from competent and incompetent participants. The simplest robust techniques will be presented below.

9.5.1 Robust statistical techniques for estimating the mean and standard deviation

The simplest robust estimate of the mean is the median. To calculate the median, all results are placed in order from low to high, and the middle point chosen (when there is an odd number of results), or the average of the two middle points when there is an even number of results.

The simplest robust technique to estimate the standard deviation is the normalized InterQuartile Range ($nIQR$). This statistic is also based on ordering the data (low to high): the first quartile (Q1) is the 25th percentile of the data and the third quartile (Q3) is the 75th percentile. There are various ways to calculate these quartiles, and different software applications might give slight differences, but all should be acceptable. The simplest manual technique is to apply the median procedure described above to each half of the ordered data (middle point if an odd number of results, average of 2 middle points if an even number). The

Interquartile range is the difference between $Q3$ and $Q1$ ($Q3 - Q1$) and then this value is “normalized” for the proportion of the IQR that would be taken by one standard deviation in a truly normal distribution, which is about 74% of the middle 50% of results. Therefore the $nIQR$ is calculated as follows:

$$nIQR = 0.7413(Q3 - Q1)$$

This is a simple and well-accepted estimate of the standard deviation that is robust to having outliers in the dataset.

9.5.2 Performance scores

The simplest performance score is available when a reference value is used with a predetermined (e.g. regulatory) limit for error. This is called the D statistic (for “Difference”) in ISO 13528:2015 (17), and can be expressed as a numeric difference or a percentage ($D\%$) depending on the expression of the allowed error. A result is evaluated as acceptable if the D (or $D\%$) is less than an error criterion, and unacceptable if the D exceeds the criterion.

The most common performance statistic for proficiency testing is the “Z score”, calculated as follows:

$$z_i = \frac{(x_i - x_{pt})}{\sigma_{pt}}$$

With

x_i = proficiency testing result i

z_i = Z score for laboratory result x_i

x_{pt} = assigned value (target value) for result

σ_{pt} = standard deviation for proficiency assessment

The Z score is commonly evaluated as follows:

$-2 \leq z \leq 2$	“Acceptable”
$-3 < z \leq -2$ or $2 \leq z < 3$	“Questionable”
$z < -3$ or $z > 3$	“Unacceptable”

There are a variety of ways to determine x_{pt} :

- as a reference value from a reference laboratory, a certified reference material or by formulation
- as a consensus value from a group of expert laboratories
- as a consensus value (robust estimate of the mean) from participants in the round of PT.

There are a variety of ways to determine σ_{pt} :

- in advance by regulation or expert judgment

- in advance based on previous rounds of PT
- using a consensus value from a group of expert laboratories
- using a consensus value from participants in the round of PT (e.g. the nIQR).

There is a wide variety of ways to calculate Z scores, so PT providers need to clearly describe the approach they are using. Similarly, laboratories need to clearly understand how their Z scores were calculated, because the interpretation could be very different for different calculations. For more information and guidance see ISO 13528:2015 (17).

9.5.3 Measurement uncertainty

Measurement uncertainty is the subject of much discussion in the laboratory-testing community. It is defined as a set of results that could reasonably be expected to occur for a given quantitative measurement of an analyte in a particular sample, under typical testing conditions. That is, if a laboratory were to test the same sample a number of times (assuming sufficient volume and stability), using the same pieces of equipment and same analysts, what would be a reasonable expectation of the range of results? Proper estimation of uncertainty can be complicated, but it can also be estimated as the statistical consensus of results from routine quality control (QC) testing, or even by the published reproducibility of a measurement method. In general, it is not common for proficiency testing providers in the medical or biological areas to request estimates of uncertainty of proficiency test results. This is discussed thoroughly in ISO 13528:2015 (17).

10 Proficiency testing round report

PT round reports should be sent to all participants to allow each participant to compare their laboratory's performance with that of peer laboratories conducting similar analysis. There are two main types of PT round reports. The first is a general preliminary report sent immediately after the closing date to allow early investigation of possible error. This report contains the expected values without individualized reporting statements. A more comprehensive individualized report is later sent to each participant. The content of this latter report can vary depending on the type of programme, however it should include:

- contact details of the EQA provider including the authorizing officer and particulars of the participant
- the type of report (interim or final), including dates
- indication that the report is confidential
- the programme type, analytes, data analysis methods and assigned values including how they were reached and the interpretation of the PT round statistics
- the individual laboratory results and recommendations depending on performance; this may be in a form of summary statistics with measure of spread, graphs and other forms of presentation.

When necessary, an amended report can be sent to participants with reference to the earlier version of the report and the reasons for sending an amended report. To emphasize the fact that the objective of EQA is to identify opportunities for improvement, the final composite report should include possible causes and suggestions for corrective actions to be acted on by relevant supervisory and management teams or authorities. The EQA provider should have established policy to indicate how the PT round report is to be used in collaboration with the EQA participants.

10.1 Response to EQA reports

It is the responsibility of the laboratory or testing site manager to share the EQA report with everyone in the laboratory or testing site. Participation in an EQA programme is an opportunity for improvement in the form of corrective action, including continuing education by writing a reflective piece. A testing site that has received a report from the organizing centre indicating a failure or unsatisfactory performance should undertake the following corrective actions:

- Review the report that was submitted to the organizing centre for clerical errors in entry of results.
- Review the report for possible transposition of specimens. For this purpose, it is essential that the laboratory retain the EQA specimens until the report is received, so that at least the internal specimen labels can be examined. In the event that specimen transposition has occurred, the laboratory should check that no patient's specimen was affected and that the labelling and verification procedure is adequate.
- If no clerical or transposition errors have been found, review the records of internal (process) quality control during the relevant time when the measurements for the PT round were made.
- Review the competence of staff undertaking the testing and their compliance with standard operating procedures. Implement additional training and competency assessment if this is a possible cause of the problem.
- Check the methodology and the instrument thoroughly to identify both human and technical failures for corrective action. This may require reference to the equipment manufacturer.
- If there is an apparent problem of commutability of results, the method, reagent or instrument may be changed; if the out-of-consensus performance is clinically significant, the laboratory might undertake a review of patients' results during the period in which the problem was evident to exclude the possibility of misdiagnosis.

NOTE Using modern telecommunications, an EQA organizing centre may organize an "on-line" PT round, whereby participants are able to compare their results immediately after measurement through communication with the centre using the Internet or e-mail. Given instant evaluation of their results, participants can immediately control their working process and take action, if required. This requires determination of target values prior to the PT round, and carries a risk of collusion among participants.

Monitoring performance over time: An EQA programme should include a procedure for monitoring performance over time. The procedure should allow participants to see the variability in their performance, whether there are general trends or inconsistencies, and whether their performance varies randomly. Graphical methods can be used to facilitate interpretation by a wider variety of readers. For some EQA programmes, longitudinal analysis can take a lot of time depending on the computing power available; therefore laboratories are encouraged to maintain their own longitudinal file.

11 EQA as an educational tool

EQA programmes generate large amounts of data that can be used by the organizers to assist in education of the participants. EQA data can improve the quality of results emanating from laboratories and other testing sites (7–9). In addition, such data can be used to convince stakeholders of the need to introduce or improve quality assurance, including internal (IQC). Methods found to have poor performance should be investigated and, if necessary, replaced by more appropriate methods, thus leading to overall improvement in the quality of results.

For participants, targeted workshops can be of great benefit in improving testing in participating laboratories. Topics which can be considered include: the importance of quality assessment, methods of quality assessment, definitions and basic statistics in quality control, interpreting internal quality control and EQA results, and resolving problems. Other topics include the pre- and post-examination factors that affect the quality of patients' results.

Continuing education of laboratory personnel and other test providers in quality management can be provided by the EQA programme organizer, professional organizations, universities and technical schools, or by the education department at the ministry of health. Although workshops may be more expensive, they are the most efficient means for educating laboratory personnel and other test providers in how to improve their performance.

Laboratory personnel and other test providers should take into account the recommendations emerging from the evaluation of PT rounds. Such information may be provided by making the PT round reports as informative as possible, and producing regular newsletters commenting on the problems encountered during the PT rounds and the technical features and interpretation of the test results. These reports should be prepared by the EQA organizer in collaboration with experts, and disseminated broadly at congresses and meetings. During meetings and workshops with participants in an EQA programme, the state of the art in laboratory practice and new developments in laboratory technology can be reviewed and observations on the performance of equipment and reagents can be communicated. It is advisable that this task be undertaken by the EQA organizer supported by experts. Government should be able to provide the funding/resources for educational programmes, and ensure that appropriate educators are made available, either from within the country or from outside. Nongovernmental organizations may also assist or contribute to the training of health workers. More difficult problems may be discussed by a group of experts at a particular laboratory or testing site.

In order to build local capacity, specialist education and training should be provided to the personnel of an EQA organizing centre in the following areas:

- laboratory safety, clinical relevance, specimen appropriateness, customer satisfaction, opportunities for improvement;
- overview of data evaluation, including the theory of measurement and statistical procedures;
- production of EQA materials, if these are prepared locally;
- critical analysis of the results of PT rounds, and corrective actions to be taken;
- concepts and use of definitive, reference and routine methods;
- national and international regulations on transport of biological materials;
- general laboratory management.

In light of the importance of education for the success of EQA, it is important to include in the programme budget the costs for training of EQA coordinating centre personnel and participants, and for new developments. Expenses for training arise mainly from the preparation of reports, workshops and visits to laboratories. In developing countries special attention should be paid to the education and training of personnel working at peripheral level, who otherwise receive little support; in these countries, costs for training should essentially be provided by the government. The suppliers of laboratory equipment and reagents can also support training courses on a voluntary basis without using the courses for advertising their products. Private laboratories usually pay for their participation through fees to the EQA programme.

The role of EQA results in accreditation: Results from EQA programmes are useful for regulatory bodies that need to evaluate the performance of laboratories covered by regulations or directives. Participation in EQA, where available, is required for accreditation (ISO 15189:2013) (12).

The use of EQA results for post-marketing surveillance: Results from EQA programmes are also useful for ministries of health and regulatory bodies to monitor the overall performance of laboratory kits, techniques and equipment, if laboratories are asked to specify the type of kits, technique or equipment they are using. If one type of kit, technique or equipment is being used by a sufficient number of participants, poor performance may be statistically ascribed to that kit, technique or piece of equipment. This finding may be used by ministries of health and regulatory bodies to change policy or, for example, to introduce stringent lot testing of reagent kits.

12 Monitoring and evaluating the programme

For an EQA programme to progress, it is important that it maintains a process of quality improvement and monitors its performance against performance indicators. Evaluation should be undertaken at least once a year and an annual report produced. Process and outcome indicators that can be used to assess the success of a programme are listed below. It should be recognized, however, that improvement in the performance of participants in relation to

outcome indicators can be influenced by factors not directly related to participation in an EQA programme, such as the introduction of improved reagents or technology.

Examples of process indicators include:

- achievement of an agreed frequency of advisory committee meetings and attendance
- achievement of the published schedule of distributions
- adequate numbers of participating laboratories and testing sites for statistically valid analysis
- proportion of laboratories and testing sites returning results for each exercise
- number of laboratory and testing sites with successful performance for each PT round
- number of problems recorded in relation to the operation of the programme
- number of complaints received regarding the operation of the programme
- number of times PT material fails to meet stability or sterility testing standards
- number of returns received noting unsatisfactory specimen quality
- achievement of published turnaround times for reports to participants
- feedback from participants
- educational meetings held
- publications by the programme.

REFERENCES

1. Joint Committee for Guides in Metrology: International vocabulary of metrology – Basic and general concepts and associated terms (VIM) 3rd edition (2008 version with minor corrections). JCGM; 2012.
2. ISO Guide 3534. Vocabulary and symbols: Parts 1–3. Geneva: International Organization for Standardization; 1993.
3. International Union of Pure and Applied Chemistry (IUPAC). Compendium of analytical nomenclature – Definitive rules (1977). Oxford: Pergamon Press; 1978.
4. ISO 9001:2015. Quality management systems – Requirements. Geneva: International Organization for Standardization; 2015.
5. Sakandé J, Nikièma A, Kabré E, Sawadogo C, Nacoulma W, Sanou M, et al. Implementation of a national external quality assessment program for medical laboratories in Burkina Faso: challenges, lessons learned, and perspectives. *Am J Clin Pathol*. 2014; 141(2):181–187.
6. Bhat V, Chavan P, Naresh C, Poladia P. The External Quality Assessment Scheme (EQAS): Experiences of a medium sized accredited laboratory. *Clin Chim Acta*. 2015; 446:61–3.
7. Sciacovelli L, Secchiero S, Zardo L, Plebani M. The role of the External Quality Assessment. *Biochimica Medica*. 2010; 20(2):160–4.
8. Kettelhut MM, Chiodini PL, Edwards H, Moody A. External quality assessment programmes raise standards: evidence from the UKNEQAS parasitology subschemes. *J Clin Pathol*. 2003; 56(12):927–32.
9. Reilly JT, Barnett D. UKNEQAS for leucocyte immunophenotyping: the first 10 years. *J Clin Pathol*. 2001; 54:7:508–11.
10. James D, Ames D, Lopez B, Still R, Simpson W, Twomey P. External quality assessment: best practice. *J Clin Pathol*. 2014; 67(8):651–5.
11. Stevens W, Gous N, Ford N, Scott LE. Feasibility of HIV point-of-care tests for resource-limited settings: challenges and solutions. *BMC Med*. 2014; 12:173.
12. ISO/IEC 17043:2010. Conformity assessment – General requirements for proficiency testing. Geneva: International Organization for Standardization; 2010.
13. ISO 15189:2013. Medical laboratories – Requirements for quality and competence. International Organization for Standardization; 2012.
14. Parekh BS, Anyanwu J, Patel H, Downer M, Kalou M, Gichimu C, et al. Dried tube specimens: a simple and cost-effective method for preparation of HIV proficiency testing panels and quality control materials for use in resource-limited settings. *J Virol Methods*. 2010; 163(2):295–300.
15. Ramos A, et al. Generation of dried tube specimen for HIV-1 viral load proficiency test panels: a cost-effective alternative for external quality assessment programs. *J Virol Methods*. 2013; 188(1–2):1–5.
16. Benzaken AS, Bazzo ML, Galban E, Pereira Pinto IC, Nogueira CL, Golfetto L, et al. External quality assurance with dried tube specimens (DTS) for point-of-care syphilis and HIV tests: experience in an indigenous populations screening programme in the Brazilian Amazon. *Sex Transm Infect*. 2014 Feb; 90(1):14–18.
17. ISO 13528:2015. Statistical methods for use in proficiency testing by interlaboratory comparison. Geneva: International Organization for Standardization; 2015.

18. World Health Organization. Guidance on regulations for the Transport of Infectious Substances 2015–2016. Geneva: World Health Organization; 2015 (http://www.who.int/ihr/publications/who_hse_ihr_2015.2/en/, accessed 30 March 2016).
19. The International Air Transport Association (IATA). Dangerous Goods Regulations. 56th edition; 2015.
20. Johnstone, A, R. Thorpe (ed.). 1987. *Immunochemistry in practice*, 2nd ed., p. 34. Oxford: Blackwell Scientific Publications; 1987.
21. Banfi G, Salvagno GL, Lippi G. Review. The role of ethylenediamine tetraacetic acid (EDTA) as in vitro anticoagulant for diagnostic purposes. *Clin Chem Lab Med*. 2007; 45(5):565–76.
22. Reardon DM, Mack D, Warren B, Hutchinson D. A whole blood control for blood count analysers, and source material for an external quality assessment scheme. *Med Lab Sci*. 1991; 48:19–26.
23. Garcia A, Subbarao S, Zhang G, Parsons L, Nkengasong J, Ou CY, et al. Impact of proficiency testing program for laboratories conducting early diagnosis of HIV-1 infection in infants in low- to middle-income countries. *J Clin Microbiol*. 2014; 52(3):773–80.
24. Nguyen S, Ramos A, Chang J, Li B, Shanmugam V, Boeras D, et al. Monitoring the quality of HIV-1 viral load testing through a proficiency testing program using dried tube specimens in resource-limited settings. *J Clin Microbiol*. 2015; 53(4):1129–36.
25. Senechal B, James VL. Ten years of external quality assessment of human immunodeficiency virus type 1 RNA quantification. *J Clin Microbiol*. 2012; 50(11):3614–9.
26. Barnett D et al. Evaluation of a novel stable whole blood quality control material for lymphocyte subset analysis: results from the UK NEQAS immune monitoring scheme. *Cytometry*. 1996; 26(3):216–22.
27. van Krieken JH, Normanno N, Blackhall F, Boone E, Botti G, Carneiro F, et al. Guideline on the requirements of external quality assessment programs in molecular pathology. *Virchows Arch*. 2013; 462(1):27–37.
28. Bancroft JD, Gamble M, editors. *Theory and Practice of Histological Techniques – Sixth Edition*. Churchill Livingstone, Elsevier; 2008.

ACKNOWLEDGEMENTS

WHO manual for organizing a national external quality assessment programme for health laboratories and other testing sites is an updated version of a previous WHO guidance titled: WHO Requirements and guidance for external quality assessment schemes for health laboratories which was published in 1999.

http://apps.who.int/iris/bitstream/10665/66089/1/WHO_DIL_LAB_99.2.pdf

It was updated by Willy Urassa with technical support from Michael Noble, Barbara De la Salle, David Bullock, Daniel Tholen, David Barnett, Sebastien Cognat, Sue Best, Mackenzie Hurlston, Mireille Kalou and Jane Carter.

The final draft was reviewed by the participants in a WHO technical working group meeting held in Geneva on 21-22 September 2015. They included: Michael Noble, Barbara De la Salle, David Bullock, Daniel Tholen, Sue Best, Mackenzie Hurlston, Mireille Kalou, Jane Carter, Stuart Scott, Sibongile Zimuto, Coumba Toure, Karen Breckenridge, Jacqueline Hall, Chonticha Kittinunvorakoon, Denise Lawrie, Sean Pdraig Mac Curtain, Karen McClure and Viktor Muchunguzi.

The WHO staff who participated in the review of the manual were: Sebastien Cognat, Anita Sands, Robyn Meurant and Irena Prat.

CONTACT:

Essential Medicines and Health Products
World Health Organization
20 Avenue Appia
CH-1211 Geneva 27
Switzerland
E –mail: diagnostics@who.int
www.who.int/diagnostics_laboratory

ISBN 978 92 4 154967 7

Prequalification Team-Diagnostics