MEETING REPORT

Guideline on the requirements of external quality assessment programs in molecular pathology

J. Han van Krieken · Nicola Normanno · Fiona Blackhall · Elke Boone ·
Gerardo Botti · Fatima Carneiro · Ilhan Celik · Fortunato Ciardiello · Ian A. Cree ·
Zandra C. Deans · Anders Edsjö · Patricia J. T. A. Groenen · Outi Kamarainen ·
Hans H. Kreipe · Marjolijn J. L. Ligtenberg · Antonio Marchetti · Samuel Murray ·
Frank J. M. Opdam · Scott D. Patterson · Simon Patton · Carmine Pinto ·
Etienne Rouleau · Ed Schuuring · Silke Sterck · Miquel Taron · Sabine Tejpar ·
Wim Timens · Erik Thunnissen · Peter M. van de Ven · Albert G. Siebers ·
Elisabeth Dequeker

Abstract Molecular pathology is an integral part of daily diagnostic pathology and used for classification of tumors, for prediction of prognosis and response to therapy, and to support treatment decisions. For these reasons, analyses in molecular pathology must be highly reliable and hence external quality assessment (EQA) programs are called for. Several EQA programs exist to which laboratories can subscribe, but they vary in scope, number of subscribers, and

J. H. van Krieken (✉) · P. J. T. A. Groenen · M. J. L. Ligtenberg ·
A. G. Siebers
Department of Pathology 824, Radboud University Nijmegen
Medical Centre, PO Box 9101, 6500 HB, Nijmegen,
The Netherlands
e-mail: J.vanKrieken@pathol.umcn.nl

N. Normanno
Cell Biology and Biotherapy Unit, INT-Fondazione Pascale,
Naples, Italy

F. Blackhall
Manchester University and Christie Hospital NHS Foundation
Trust, Manchester, UK

E. Boone
Laboratory for Molecular Diagnostics, H. Hartziekenhuis,
Roeselare-Menen, Roeselare, Belgium

G. Botti
Dipartimento di Anatomia Patologica e Citopatologia, IRCCS
Istituto Nazionale Tumori, Fondazione G. Pascale, Naples, Italy

F. Carneiro
Institute of Molecular Pathology and Immunology of the
University of Porto (IPATIMUP) and Medical Faculty/ Centro
Hospitalar de São João, Porto, Portugal

I. Celik
Merck Serono Global Research & Early Development—Oncology
Platform, Merck KGaA, Darmstadt, Germany

F. Ciardiello
Dipartimento Medico-Chirurgico di Internistica Clinica e Sperimentale
F. Magrassi e A. Lanzara, Seconda Università di Napoli,
Naples, Italy

I. A. Cree
Warwick Medical School, Clinical Sciences Building, University
Hospitals Coventry and Warwickshire, Walsgrave,
Coventry, UK

Z. C. Deans
UK NEQAS for Molecular Genetics, UK NEQAS, The Royal
Infirmary of Edinburgh, Edinburgh, UK

A. Edsjö
Clinical Molecular Pathology Unit, Clinical Pathology and
Cytology, Sahlgrenska University Hospital and Sahlgrenska
Cancer Center, the Sahlgrenska Academy, University of
Gothenburg, Gothenburg, Sweden

O. Kamarainen
European Molecular Genetics Quality Network (EMQN) Genetic
Medicine, St. Mary's Hospital, Manchester, UK

H. H. Kreipe
Institute of Pathology, Hannover Medical School, Hannover,
Germany

M. J. L. Ligtenberg
Department of Human Genetics, Radboud University Nijmegen
Medical Centre, Nijmegen, The Netherlands

execution. The guideline presented in this paper has been developed with the purpose to harmonize EQA in molecular pathology. It presents recommendations on how an EQA program should be organized, provides criteria for a reference laboratory, proposes requirements for EQA test samples, and defines the number of samples needed for an EQA program. Furthermore, a system for scoring of the results is proposed as well as measures to be taken for poorly performing laboratories. Proposals are made regarding the content requirements of an EQA report and how its results should be communicated. Finally, the need for an EQA database and a participant manual are elaborated. It is the intention of this guideline to improve EQA for molecular pathology in order to provide more reliable molecular analyses as well as optimal information regarding patient selection for treatment.

Keywords Molecular pathology · External quality assessment · Oncology · Guideline

Introduction

Molecular pathology, defined as the analysis of nucleic acids in tissue or cell samples, has become an integral part of daily diagnostic pathology. In oncology, it is used to characterize or classify tumors, to detect specific molecular alterations that relate to prognosis, or define targets that predict therapy response and other treatment decisions. The results of these tests directly influence management of individual patients, which is referred to as personalized therapy [1]. Recently developed drugs, which target molecular pathways altered in cancer cells, are only effective in the subset of patients that carries a molecular alteration that is targeted by a particular drug. For the identification of such molecular alterations the term "predictive molecular pathology" has become popular. An example of personalized therapy is the use of *EGFR* gene mutation analysis prior to prescription of *EGFR*-targeting drugs in patients with non-small-cell lung carcinoma [2, 3]. The requirements for the reliability of molecular pathology are high since the results, which generally extend beyond histologically recognizable subtypes, are used to determine the eligibility of a patient for treatment using a specific class of drug and unreliable results might lead to over- or undertreatment of patients. Since these drugs are expensive, the availability of reliable tests will also significantly improve the cost effectiveness of these new treatment modalities.

In view of their widespread use in clinical practice, molecular tests need to be both accurate and readily available. The challenge is to ensure that a sufficient number of laboratories can provide reliable test results. Contrary to the USA, where in vitro diagnostic product (IVD) regulation has been developed, in Europe no regulatory framework exists on which assay(s) are eligible as drug response marker. Many different testing methods exist, which under appropriate laboratory and expertise conditions might provide reliable results but the equivalence of the results can be only

A. Marchetti
Center of Predictive Molecular Medicine, University-Foundation, Chieti, Italy

S. Murray
BioMarker Solutions Ltd., London UK & GeneKor A.E., Athens, Greece

F. J. M. Opdam
Molecular Diagnostics; Quality Control Business, Life Technologies Europe BV, Bleiswijk, The Netherlands

S. D. Patterson
Medical Sciences Amgen Inc., Thousand Oaks, CA, USA

S. Patton
European Molecular Genetics Quality Network (EMQN) Genetic Medicine, St. Mary's Hospital, Manchester, UK

C. Pinto
Medical Oncology, S.Orsola-Malpighi Hospital, Bologna, Italy

E. Rouleau
Institut Curie, Paris, France

E. Schuuring
Department of Pathology and Medical Biology, University Hospital Groningen and University of Groningen, Groningen, The Netherlands

S. Sterck · E. Dequeker
Department of Public Health, Biomedical Quality Assurance Research Unit, University of Leuven, Leuven, Belgium

M. Taron
Medical Oncology Service, Catalan Institute of Oncology, Hospital Germans TriasiPujol, Badalona, Spain

S. Tejpar
Department of Oncology, Laboratory of Molecular Digestive Oncology, University of Leuven, Leuven, Belgium

W. Timens
Department of Pathology and Medical Biology, University of Groningen and University Medical Center Groningen, Groningen, The Netherlands

E. Thunnissen
Pathology, VU University Medical Center, Amsterdam, The Netherlands

P. M. van de Ven
Department of Epidemiology and Biostatistics, VU University Medical Center, Amsterdam, The Netherlands

M. Taron
Pangaea Biotech-Dexeus Institute, Barcelona, Spain

established through inter-laboratory comparison. To attain this goal, external quality assessment (EQA) programs are essential. The notion that the reliability of molecular tests depends on such programs was soon shared by many groups, and presently, there are several accessible EQA programs for the assessment of the reliability of gene aberration testing (e.g., http://kras.eqascheme.org; http://www.ukneqas.org.uk; http://www.emqn.org/emqn/schemes). The results of these programs clearly indicate the need for EQA, as some laboratories participating in these activities provided results below the standards set by the EQA provider [4–6]. Furthermore, as yet there is no harmonization of the standards applied by these programs, which often vary in scope, number of subscribers, and execution. Therefore, an expert group of clinical oncologists, pathologists, molecular biologists, EQA providers, and representatives from the pharmaceutical industry agreed to develop guidelines for EQA in molecular pathology. In a first meeting during the European Congress of Pathology in Helsinki in September 2011, an outline of a guideline document was defined. A draft text was circulated, and the document was discussed during a meeting hosted by the Italian Association of Medical Oncology in Naples in 2012. The meeting was also endorsed by the European Society of Pathology, the European Society of Medical Oncology, and the Italian Society of Pathology and Cytopathology. The document has not been formally approved by these societies since this process is time consuming as it involves consultation of the governing bodies of these professional organizations as well as their membership. Therefore, these guidelines represent a document conceived by professionals that have been identified by these organizations as experts on this specific topic. The focus of the meetings was on tests based on DNA extracted from cell or tissue samples, as they are considered fundamentally different from tissue-based techniques with a microscopy read-out such as immunohistochemistry and (fluorescent) in situ hybridization. The general scope of the document is applicable to all types of molecular pathology testing, even though some aspects of the guideline (e.g., the used marking criteria) are only relevant for some tests. The document intends to harmonize EQA in molecular pathology and to ensure that EQA programs reflect the diagnostic and clinical reality as closely as possible, assuring users of these services that the laboratories performing these tests provide results according to accepted predefined standards of quality.

Organization of an EQA program

The development and operation of EQA programs shall be undertaken by providers with competence to conduct inter-laboratory comparisons and access to the required techniques and samples. For EQA programs, it is considered of primary importance to use samples that reflect the diagnostic setting as much as possible. The large majority of molecular tests in pathology are performed on formalin-fixed paraffin-embedded (FFPE) tumor tissue, and an EQA program needs to reflect this, by using either FFPE-tumor tissue samples or close mimics thereof.

An EQA program is oriented towards a specific application, mostly the detection of a specific alteration in a specific tumor type. Therefore, EQA programs in molecular pathology need to be developed by groups of content experts. Expert groups should include at least medical experts, experts on the relevant laboratory methods, and representatives of the EQA provider. The medical expert(s) should include pathologists with proven experience in molecular techniques and oncologists with proven experience in evaluating molecular alterations in the appropriate morphological context and both with knowledge of the clinical and pathological background in the domain of the scope of the EQA program. The method expert(s) should have experience in methods of molecular analysis and knowledge of the molecular context and of the technologies used for diagnostic testing. The EQA provider is responsible for the organization and management of the EQA program in accordance with ISO 17043 [7]. Therefore, the provider needs to have experience in quality management, a solid background in the diagnostic domain of the EQA program, and the necessary facilities to run such a program. For each EQA program, the EQA provider should appoint an EQA coordinator. The EQA coordinator is responsible (together with the provider and the medical and technical expert(s)) for the selection and distribution of samples, reception of the results, analysis of results, reporting to participants, and to regulatory or certifying agencies as required. The composition of an EQA program should fulfill the minimal requirements as proposed in this guideline (see below).

Three aspects of molecular testing should be covered by an EQA program:

- Pathology review. Considering the diversity and heterogeneity of tumor tissue, pathology review and assessment of section quality is mandatory. For instance, due to limitations of most routinely utilized techniques, it is important to determine the percentage of neoplastic cell content in the material to be analyzed.
- The molecular test itself, including DNA extraction, validation of the methods used for the test, and accuracy of the result.
- Reporting with particular attention to the following aspects: (1) identification of the sample(s) analyzed, (2) information on the type of assay used, (3) adequacy of the sample relative to the underlying request and the test used, and (4) accurate assessment of the clinical implications of the result.

    The best practice is to assess the pathology review, the molecular analysis, as well as reporting of the results

in one EQA program. Given the fact that laboratories need to participate on regular basis in EQA programs, it is advisable for the EQA program to be offered at least once a year.

Setup of an EQA program assessing the entire workflow

Ideally, an EQA program distributes the same samples to all participating laboratories to allow inter-laboratory comparison of the same starting material [8]. However, when a large number of laboratories subscribe the samples are often provided by reference laboratories that prepare EQA material for a smaller number of laboratories. This will result in different participants testing different samples, in which case the program should be closely regulated and controlled by the EQA coordinator who undertakes the central evaluation of the results (in collaboration with the medical and technical experts) and ultimately the overall responsibility for the program.

The participants are expected to provide information using a standard EQA questionnaire on the analyte(s) tested for, the methods used (including major test characteristics, e.g., analytical sensitivity (and in exceptional cases raw data), the percentage of neoplastic cells present, and the obtained result(s).

All the EQA submissions are required to be evaluated independently by at least two assessors under the responsibility of the EQA provider and the medical and technical expert(s).

Sample selection

The selection of material that is used in an EQA program represents an important issue. In addition to routinely processed surgical pathology material, synthetic samples composed of well-defined cell lines with known mutation analyte status and copy number of target genes (if required) can be used. The former offers the important advantage that the complete analytical pathway is assessed; the latter that the proportion of cells with the defined aberration(s) can be strictly defined and that homogenous material can be distributed to a larger number of laboratories.

The set of validated samples that should be distributed to the participating laboratories should reflect the range of clinically relevant molecular alterations present in the clinical samples. When the volume of the starting material is limited, it might be impossible to send all participants sections derived from the same tissue specimen.

The selection of samples for EQA programs has to take into account two different intentions. One is to mirror the daily diagnostic situation and to use samples that contain the most common aberrations. The other is to enrich for difficult cases and those with sample characteristics on the limit of the acceptable which might uncover latent weaknesses in test

performance and interpretation of results [9]. When the latter approach is followed, the number of laboratories that fail to perform according to the preset standards might be higher. However, a laboratory of which the analytical sensitivity is adequate for samples with a fraction of neoplastic cells above the threshold, should not be scored as insufficient when it fails to identify the aberration or analyte in a sample with a fraction of neoplastic cells or analyte below the threshold. In this way, a misleading representation of the performance status of a molecular pathology laboratory can be avoided. Samples that do not reflect clinical practice should not be used. EQA providers should be aware that in selecting a balanced set of difficult and mainstream samples, a potential bias in the performance status of participating laboratories can be avoided. The samples provided to participants should be sections of paraffin-embedded material from different tissue specimens harboring a possible aberration. The amount of material supplied should reflect the clinical situation; this is to be determined by the EQA provider together with the medical and technical expert. The EQA provider should supply the participants with information on fixation and integrity of the samples. Clinical information and the request for testing (treating physician or other responsible) should be supplied for those samples for which the EQA provider requests the delivery of a full report. The materials distributed are provided as specimens for the sole purpose of enabling external quality assessment for the aberration stated at the recipient's laboratory for that particular distribution.

Turnaround time

The EQA provider shall define a turnaround time that reflects the clinical situation. Normally, a turnaround time of 10 working days after receiving the samples is considered reasonable for EQA samples. This prevents an approach for such samples from routine procedures. Results that are received after the reporting deadline should not be accepted, unless an explanation is given, deemed satisfactory by the EQA provider.

Report

Laboratories are expected to submit a report, comparable to what is normally generated by the diagnostic service for a specific sample type. These reports should be evaluated according to a guideline on reporting of molecular testing (a detailed guideline on reporting will be published separately).

Evaluation

The EQA coordinator should coordinate the evaluation of the submitted results and raw data in close collaboration with the program organizers. The results should be assessed independently by professionals with (extensive) experience in

the field of diagnostic molecular pathology, by comparing them with validated results and using predefined criteria. The medical and technical expert(s) of the EQA program should be involved in the assessment process. The result(s) and interpretation of every report should be independently evaluated by at least two members of the assessment team. Results should subsequently be discussed during an assessment meeting. The final scores and program feedback comments should be reviewed by the medical expert(s), technical expert(s), and EQA provider, in order to provide consistent assessment across participants and across molecular pathology EQA programs.

Results of the EQA programs should be distributed after having been approved by the assessors as well as the medical and technical expert(s) and EQA provider. The results should be made available anonymously to all participants and each participating laboratory should receive individual feedback. All participants should receive a certificate of participation after submission of their data. In addition, each laboratory performing according to preset standards should receive a certificate of performance. Participants with a score above the predefined threshold (see "Numbers of samples needed for an EQA scheme") might be listed on the website of the EQA program if the program supports this.

Criteria for a reference laboratory

Selection

For the preparation and validation of the samples, reference laboratories are required. These laboratories are selected on three criteria [5]: (a) proven experience with the diagnostic test(s) comprised in the EQA program, (b) adequate supply of samples (blocks) that can be used for the EQA program (according to the national legal requirements of the use of patient samples), and (c) ability to coordinate and execute EQA programs in collaboration with the EQA provider if required. Ideally, for every program, the reference laboratories cover the techniques that are suitable for aberration detection in terms of sensitivity and specificity and that are most commonly used in routine practice in order to avoid any methodological bias.

Minimum requirements

The reference laboratory should be a fully equipped molecular pathology laboratory with certified pathologists with proven experience in molecular pathology, clinical molecular biologists, and technicians. The laboratory should be accredited to a recognized international standard (e.g., ISO 15189 [10]) and should have passed an EQA test.

Validation of EQA samples

The EQA provider and the medical and technical experts select the program samples. Selected samples need to have been pretested in at least two reference laboratories and their findings need to be the identical and will serve as the standard. Only samples which are tested by all reference laboratories with identical results should be entered into the EQA scheme [4–6, 11, 12]. The laboratories that prepare the samples should send the last tissue section of the series cut for distribution to a central reference laboratory for independent validation.

Requirements for EQA test samples

This guideline focuses on EQA programs for DNA-based tests on FFPE specimens, since these are presently by far the most commonly used samples for molecular testing in pathology. For other sample types, different requirements may be applicable.

Cell lines

Artificial FFPE tissue blocks can be generated by paraffin embedding of a formalin-fixed mixture of cell lines with and without specific aberrations. The used cell lines need to be genetically stable and to contain a known copy number of the genes of interest. Cell lines need to be banked for future use in the EQA scheme in view of its repetitive nature. Batch validation ensures their acceptability as control material. Control material with a CE marked for IVD use is preferable as this is homogeneous in composition and the number of potentially available samples is not limited. They have the disadvantage of not reflecting the complex tissue composition of human tumors and issues related to tumor heterogeneity.

Tissues

Theoretically, human tumor tissue constitutes the preferred control sample. Its availability is, however, limited by the amount of available human tissue and by issues related to transport of human tissue samples across national borders. There is also additional work and time needed for the generation of suitable materials for EQA programs, such as validation of the specimen both by pathology review and molecular analysis of at least the first and last cut section to exclude absence of the aberration due to differences in the proportion of tumor tissue in a sample or tumor heterogeneity.

Informed consent

Informed consent is not a mandatory prerequisite for the use of any patient derived material, since samples for test

validation are exempt from research regulations requiring informed consent. The EQA provider, including laboratories which provide material, should be diligent in ensuring that activities associated with specimen collection are conducted in compliance with the appropriate legal and regulatory requirements, in particular regarding privacy protection issues.

Percentage of neoplastic cells

The percentage of neoplastic cells in the material from which the analyte (e.g., DNA) is isolated is crucial for the use of patient samples for molecular testing and for a laboratory to determine whether they consider the sample sufficient for analysis. Due to limitations in the analytical sensitivity of several techniques routinely utilized in molecular pathology, it is important that this percentage is included in the report of the analysis. In the event of test samples being provided in an Eppendorf tube rather than as a slide, the percentage of neoplastic cells needs to be predefined and supplied to the participant or a separate mounted section for estimation of neoplastic cell content by (virtual) microscopy needs to be supplied.

Quality and amount of the analyte

In general, the quality and amount of the analyte (e.g., DNA) present in the provided samples should be consistent with the different technologies participating laboratories are using in routinely diagnostic practice.

Quality (e.g., DNA)

For control materials derived from surrogate sources and fixed using routine procedures, the quality of the DNA should theoretically be similar to that of DNA isolated from FFPE material.

For patient material, depending on the source (e.g., FFPE or bronchio-alveolar lavage), the quality of the analyte (e.g., DNA) will be highly heterogeneous and this should be reflected in the samples sent out where possible. However, prior analysis of all samples by the reference laboratories should guarantee that the quality of the analyte (e.g., DNA) is sufficient for the requested molecular analysis.

Quantity (e.g., DNA)

This is directly related to the quality of the analyte. For FFPE specimens in general, between 10–50 ng of DNA is sufficient to perform a single PCR-based analysis. However, several IVD and CE marked kits for gene mutation detection specify a higher minimum DNA quantity. Moreover, dependent on the required target and methodology used, multiple

analyses may need to be performed in parallel. It is important that EQA schemes adhere to the upper limit of the available techniques, thus allowing participation of all laboratories in the program.

Number of samples needed for an EQA scheme

An important issue is the number of samples that are used in EQA programs. Currently, different approaches exist with respect to the result characteristics and number of EQA samples which are to be tested. Reliable evaluation needs to be based on at least 10 samples that may be analyzed in one batch or in different smaller batches that are sent within a year.

Both small and larger sample sets are currently used in different EQA programs. In the small samples set, a limited number of cases (e.g., 3) are provided but with frequent distributions (e.g., 3 different EQA rounds per annum), and in the larger set, more samples (e.g., 10) are distributed but only once per year. The larger set offers the opportunity to use confidence limits around the outcome of the EQA [13], though this is also possible over time using smaller sample numbers, which are sent out more frequently.

In theory, a 90 % credible interval (CI=Bayesian confidence interval) may be calculated for a number of samples in external quality assessment (see Table 1). For instance, in a lab with 8 out of 10 correct answers in an EQA study, the chance is 90 % that between 53 and 92 % of the routine analysis will be correct. There is a 5 % chance that this lab generates the correct answer in more than 92 % of the cases in routine diagnostic practice. As it seems reasonable to set a norm in external quality assessment at the smallest number of correct cases where the upper limit of the 90 % CI is above the 95 %, theoretically in a setting of 10 cases at least 9 cases should be reported correctly. This approach offers a [generous] limit for poor performance: upper limit below the 95 % level.

To provide a certificate for good performance with statistical support at the same 5 % level, with the lower 90 % CI limit >95 % would require 58 samples all to be tested correctly (90 % CI [95.0, 100]). For 92 samples, a correct test of 91 would result in a 90 % CI (95.0, 99.6). Such a number of cases in an EQA program only to statistical underscore good performance is not realistic. Therefore, we propose a two-tiered approach: establishing adequate and poor performance, with the dividing line between the two defined as the upper boundary of the 90 % confidence limit.

For a single EQA round with a smaller sample size, e.g., 3 samples, statistical underscoring is not feasible. In this case, a substandard level of performance is judged in case of any critical error, i.e., performance should be at a 100% level (3 out of 3 correct) instead of 67 % level (2 out of 3 correct).

Table 1 For a specific set of samples in external quality assessment (n=10, 14, 20, and 30), the 90 % credible interval (CI, the region between the 5th percentile and 95th percentile of the posterior probability distribution) is shown for the success rate (fraction of correct answers×100 %)

N samples	10	10	14	20	30	30
Number of correct answers	Number of correct answers	90 % CI	90 % CI	90 % CI	Number of correct answers	90 % CI
n/n	10/10	76.2–99.5	81.9–99.7	86.7–99.8	30/30	90.8–99.8
$n-1/n$	9/10	63.6–96.7	72.1–97.6	79.3–98.3	29/30	85.6–98.8
$n-2/n$	8/10	53.0–92.1	63.7–94.3	72.9–96.0	28/30	81.1–97.3
$n-3/n$	7/10	43.6–86.5	56.0–90.3	67.1–93.2	27/30	76.8–95.5
$n-4/n$	6/10	35.0–80.0	48.9–85.8	61.6–90.1	26/30	72.9–93.4
$n-5/n$	5/10	27.1–72.9	42.3–80.9	56.3–86.8	25/30	69.0–91.2
$n-6/n$	4/10	20.0–65.0	36.0–75.6	51.3–83.2	24/30	65.3–88.9
$n-7/n$	3/10	13.5–56.4	30.0–70.0	46.4–79.4	23/30	61.7–86.5
$n-8/n$	2/10	7.9–47.0	24.4–64.0	41.7–75.5	22/30	58.2–83.9
$n-9/n$	1/10	3.3–36.4	19.1–57.7	37.2–71.4	21/30	54.8–81.3
$n-10/n$	0/10	0.5–23.8	14.2–51.1	32.8–67.2	20/30	51.5–78.7

The 90 % CI is constructed with Bayesian statistics [14] assuming uniform prior probability for the success rate on the interval between 0 and 1

In case of a 10 sample round, of which 9 need to be correct, not all samples have to be assessed in one batch. Combining the results of three test rounds within a year, of three to four samples each, will provide the same statistical confidence but only if the samples in the set are all different. Assessing samples over time has the added advantage of a more continuous assessment of quality.

Scoring system for EQA in molecular pathology

EQA participant results and reports should be assessed against predefined peer reviewed criteria in a scoring system. In a scoring system for EQA in molecular pathology, the following distinct categories have to be distinguished:

- The "pre-analytical" phase, which includes examination of the sample by a pathologist, assessment of the adequacy of the test sample, evaluation of the percentage of neoplastic cells and whether or not the sample needs to be dissected
- The "analytical" phase, i.e., DNA isolation and genotyping
- The "post-analytical" phase including interpretation and reporting of the results of the analysis.

Scoring of the pre-analytical phase

The items to be marked in this category are assessment of the adequacy of the sample, evaluation of the percentage of neoplastic cells, and of the need for dissection. This phase can also include other parameters related to the histological evaluation of the sample. However, quality assurance of histopathological diagnosis of the specimens does not fall within the scope of this guideline. Earlier EQA schemes on *KRAS*

mutation testing in colorectal cancer led to the conclusion that a gold standard for estimating the percentage of neoplastic cells does not exist [2]. As this is a crucial element in the pre-analytical phase, guidelines to achieve improved results for this issue are currently being developed but these can only be applied after having been formally validated by the stakeholders. This guideline does not recommend to scoring the pre-analytical phase, although such data might be collected for documentation and for training purposes, as records of pre-analytical performance might be useful in advising a laboratory that is confronted with discordant results.

Scoring of the analytical phase

Genotyping is the core of most currently practiced tests in molecular pathology and consequently also its EQA. A detailed scoring system for genotyping performance is mandatory.

Although different scoring criteria can be considered, this guideline proposes to use a standard approach for all EQA programs which will facilitate cross-comparison. The proposed evaluation system is based on a score of two points for each correctly tested sample, with minus points in case of error and in number depending on the type of error (see Table 2).Assessment should be performed by a board of experts familiar with the problems that may be encountered with the range of methods applied by the participating laboratories. Table 2 lists the most common errors described in EQA programs and provides an example as to how the evaluation might be performed. Finally, evaluation of the results must also take into account which method has been used for the test. For example, some Real-time PCR-based methods do not distinguish between different mutations in the same codon. Likewise, for some commercially available kits, the validation studies limit the result to whether a

Table 2 Proposed evaluation system with aberration scores for the most common alternatives encountered in EQA programs

Marking criteria	Marks
Aberration correctly identified	2.00
Aberrationnot correctly identified (wrong aberration, false positive, or false negative)	0.00
Aberration miss-positioned or miss-called (e.g., incorrect base/amino acid detected for genotypes)	1.00
Error in aberration nomenclature which could be misinterpreted	1.50
Small error in aberration nomenclature with no impact	Comment
Not correctly using HGVS nomenclature for either the nucleotide or amino acid changes (marks to be deducted once only)[a]	1.50
Test failure giving no result for the sample	0.50

HGVS Human Genome Variation Society (http://www.hgvs.org/)

[a] Unless the test result only designates mutant or wild type (then specific alleles to which that designation applies should be included)

sample is mutant or not. Such methods are not to be preferred but formally a reduced score cannot be given in such a case because of intrinsic limitations of the chosen method. If 10 samples are used in an EQA, the maximum score will be 20. An average score can be calculated by dividing the final score by the number of samples.

Scoring of the post-analytical phase

An EQA program includes a description of mock clinical cases for each (or a selected) sample and requires the submission of a formal clinical report in which the available clinical information is taken into account. These reports are also scored and should reflect the reports as they are issued in daily practice to requesting physicians. Several aspects of interpretation of the test results and editing of the report can be marked. Before dispatching the sample, the assessors should determine which key elements of interpretation should be present in the reports based on expert consensus and best practice guidelines. The proposed scoring system should be applied to all cases for which a report is requested within a program.

The following themes should be scored/commented as follows:

1. Patient identification: Table 3
2. Report look and content: Table 4
3. Interpretation (includes both the biological interpretation of the result and the clinical interpretation, i.e., the suggestions for adequate clinical management): the total score is 2.00 points. The assessors should predefine

which elements will be scored regarding interpretation and content of the written report (Table 5).

A laboratory should receive marks for each appropriate interpretation element in the report, and no marks should be awarded in case of an absent or incorrect element. The sum of the marks (further referred to as the "interpretation score") is a maximum 2.00 per case.

Finally, scoring of the different elements—result, patient identification/clerical accuracy, and interpretation—should be reported separately in the EQA final report The pass/fail level for each case should be predefined by the EQA provider.

Consequences of poor performance

In many European countries, participation in EQA programs is becoming part of daily practice, but standards for EQA and definitions of performance levels are lacking. In the USA, the "Clinical Laboratory Improvement Act" of 1988 defines *unsatisfactory* performance in EQA as failure to attain the minimum satisfactory score for an analyte for a single testing event (http://wwwn.cdc.gov/clia/regs/subpart_h.aspx).

Table 4 Proposed marking of report look and content

Marking criteria	Marks
Length of report more than one page	Comment
Spelling and typographic error (excluding patient identifiers)	Comment
Errors in sample identifiers (arrival date of the sample, sample number)	Comment
Failure to provide a clear presentation of results	Comment
Failure to describe limitations of tests	Comment
Percentage of neoplastic cells present within the sample not listed	Comment
Lack of name/address referral person	Comment
Identification of report authorizer	Comment

Table 3 Proposed scoring of patient identification

Marking criteria	Marks
Correct name and first name of the patient without clerical errors	1.00
Date of birth without any error	1.00

Table 5 Proposed scoring for items of biological and clinical interpretation

Marking criteria	Marks
Result and explanation how to interpret is given	0.75
Clinical interpretation is given (e.g., prediction of effect of genotype on therapy response)	0.50
Reference sequence and version correctly used throughout the report	0.25
Specification of the molecular assay employed, and the aberration(s) it detects (or regions/exons covered) are stated	0.25
Limitations of the assay/sensitivity of the test is stated	0.25

Unsuccessful performance is defined as failure to attain the minimum satisfactory score for either two consecutive or two out of three consecutive events. Failure to return results to the EQA provider within the time frame specified by the program or failure to participate is also classified as unsatisfactory performance. In clinical chemistry and microbiology in the USA, standard practice is to test samples for a greater number of times. Failure to attain an overall testing event score of at least 80 % is classified as unsatisfactory performance. It is the responsibility of the labs performing the molecular test to take measures towards improvement, following an "unsatisfactory" performance test result. Essential to attain this goal is regular internal quality control and result validation using externally provided samples. In the absence of the mandatory implementation of laboratory accreditation across Europe, the consequences of "unsuccessful" performance in molecular diagnostics in Europe have not yet been determined. One option is that the insufficiently performing laboratory withdraws the test from its test catalogue. Alternatively, either professional organizations or the government should determine what the consequences will be. It is not the responsibility of the EQA organizer to impose sanctions but it can provide help and support to the contested laboratory in order to improve its service. This might involve providing reference material, methodological advice, or support in quality management through reviewing plans for correction and prevention. For the laboratories that attain a sufficient score, it is recommended that they be listed on the website of the EQA provider.

Content of an EQA report/communication of EQA results

A general program report and scores of individual laboratories including specific comments should be published, once the program has been completed. It is strongly recommended that the results of the EQA testing be communicated to the participants before a general report is emitted, for instance immediately after submission of the results. In this way, laboratories can instantaneously review their performance scores and take, in case of need, appropriate measures to improve their practice.

The participants will receive a certificate of participation and a certificate of performance. In addition, two reports should be provided to each participant, one general report summarizing the anonymized results of all participants and one that is specific for each participating laboratory and provides a performance appraisal with individualized comments and feedback. These reports should both review performance and provide an educational component including comments regarding performance of the group as a whole.

The reports can be provided as hard copies or electronically through a password protected website or by email. The EQA provider should follow procedures to prevent unauthorized access or amendment of these reports. Any comments interpreting the EQA finding should be explanatory and without ambiguities. Reports to participants should be validated prior to dispatch and should be sent to participants in a timely manner with the dispatch date recorded for audit purposes. Participants should have an opportunity to appeal the received reports, after which the final general report can be released.

General report

Unless it is not applicable or the EQA provider has valid reasons for not doing so, reports should include the following elements (ISO 17043 [7]):

1. The name and contact details of the EQA provider;
2. The name and contact details of the coordinator of the EQA program;
3. The name(s), function(s), and signature(s) or equivalent identification of person(s) authorizing the report;
4. An indication of which activities are subcontracted by the EQA provider;
5. The date of issue and status (e.g., preliminary, interim, or final) of the report;
6. Page numbers and a clear indication of the end of the report;
7. A statement on the extent to which results are confidential;
8. The report number and clear identification of the EQA program;
9. A clear description of the EQA items used, including necessary details of the EQA;
10. Description of the preparation and the homogeneity and stability assessment of the EQA sample(s);
11. The participants' results, both individual and aggregate group results;
12. Statistical data and summaries, including assigned values and range of acceptable results and graphical displays;

13. Procedures used to establish any assigned value;
14. Details of the metrological traceability and measurement uncertainty of any assigned value (where applicable);
15. Procedures used to establish the standard deviation for proficiency assessment, or other criteria for evaluation;
16. Assigned values and summary statistics for test methods/procedures used by each group of participants (if different methods are used by different groups of participants);
17. Comments on participants' performance by the EQA provider and technical advisers;
18. Information about the design and implementation of the EQA program;
19. Procedures used to statistically analyze the data;
20. Advice on the interpretation of the statistical analysis;
21. Comments or recommendations, based on the outcomes of the EQA program;
22. Details and identity of the cell lines used to generate EQA panel samples and the results of the reference laboratories on these samples;
23. Details of the patient specimens included in the panel and the results of reference laboratories on these samples.

Individual comments to the participants

Where appropriate for the purpose of the EQA program, the EQA provider should offer expert comments on the overall performance in comparison with prior expectations, taking into account;

- Variation within and between participants, and comparisons with any previous external quality assessment rounds, similar proficiency testing programs, or published precision data;
- Variation between methods or procedures;
- Possible sources of error (with reference to outliers) and suggestions for improving performance;
- Advice and educational feedback to participants as part of the continuous improvement of participants procedures;
- Situations where unusual factors make evaluation of results and comments on performance impossible;
- Any other suggestions, recommendations or general comments;
- Conclusions.

Additional reporting responsibilities

When required by law, the EQA provider may have to report results not only to the participants, but also to regulatory agencies.

EQA databases

Medicine, and especially clinical oncology, is in a phase of rapid change, thanks to the development of targeted therapies. These new types of therapy form the basis for the present molecular tests that require EQA. The multidisciplinary group felt it as a joint responsibility that the data generated by the many laboratories that participate in EQA should be combined by linking the databases. This should enable a rapid growth of knowledge on the general performance level and expertise of participating laboratories. A subgroup of interested specialists from different disciplines currently investigates the possibilities to create such a database. The European Society of Pathology intends to develop a database in which specific results of all EQA providers can be collected to obtain a better overall view on how performance testing in molecular pathology evolves. Furthermore, EQA providers should be encouraged to make the results of their programs available to the public domain.

Participants manual

The EQA provider should give participants sufficient prior notice before sending EQA samples, in the form of a participants' manual. General information about the program should be provided through a website, flyer, or catalogue.

The instructions to the participants should include:

- General information about the EQA program such as the aims of the EQA program, and the frequency of the EQA program (e.g., number of samples and events per year)
- The EQA process and practical details of the program design and procedures
- The terms and conditions of participation, organizational structure of the program, and responsibilities of the different parties
- Details of how to apply for participation.

Each participating laboratory should be assigned a unique EQA code/identification number. This should be used in all correspondence between the laboratory and the program administration. This code number should not be disclosed by the EQA provider to any third party without prior written permission from the primary laboratory contact.

The participating laboratory must provide a primary contact (e.g., Head of Laboratory, Quality Manager) responsible for registering the laboratory with the EQA program provider and to act as the link between the EQA provider and the laboratory. It is the responsibility of the laboratory to keep the contact information to the EQA provider up to date.

The laboratory should participate in available EQA program in such a way that all the tests the laboratory performs as

a clinical service are covered. In many countries, this is now a requirement for laboratory accreditation (ISO: 15189 [10]).

EQA samples must be handled in exactly the same way as clinical samples. If this is not applicable because of the use of non-routine material for the EQA sample (such as cell lines or extracted DNA), handling of the sample should be as close as possible to that of daily diagnostic specimens.

The program organizer has to make clear that materials distributed are provided as specimens for the sole purpose of enabling external quality assessment at the recipient's laboratory.

The name of the laboratory and the assessment of individual laboratory performance are confidential to the participant and should not be released by program organizers without the written permission of the head of the laboratory to any third party.

The EQA provider should be committed to providing quality assessment for all molecular pathology testing laboratories which fulfill the requirements of a participating laboratory.

Concluding remarks

Even though the use of molecular tests for selection of patients for targeted therapy is relatively new, there is already a large body of evidence that EQA programs for laboratories performing such tests are needed and can improve laboratory performance. During the multidisciplinary meeting in March 2012 in Naples, many issues regarding EQA were raised and in most cases consensus could be reached. This consensus is described in the present report. Several issues were not entirely resolved and more experience is needed to further improve EQA. For now, we hope and expect that this guideline will lead to improved EQA for molecular pathology and as a result to better treatment selection for patients.

References

1. Dancey JE, Bedard PL, Onetto N, Hudson TJ (2012) The genetic basis for cancer treatment decisions. Cell 148:409–420
2. Pao W, Chmielecki J (2010) Rational, biologically based treatment of EGFR-mutant non-small-cell lung cancer. Nat Rev Cancer 10:760–774
3. Penzel R et al (2011) EGFR mutation detection in NSCLC—assessment of diagnostic application and recommendations of the German Panel for Mutation Testing in NSCLC. Virchows Arch 458:95–98
4. Dequeker E, Ligtenberg MJL, Vander Borght S, Van Krieken JJM (2011) Mutation analysis of KRAS prior to targeted therapy in colorectal cancer: development and evaluation of quality by a European external quality assessment scheme. Virchows Arch 459:155–160
5. Bellon E et al (2011) External quality assessment for KRAS testing is needed: setup of a European program and report of the first joined regional quality assessment rounds. Oncologist 16:467–478
6. Wong NACS, Deans ZC, Ramsden SC (2012) The UK NEQAS for molecular genetics scheme for gastrointestinal stromal tumour: findings and recommendations following four rounds of circulation. J ClinPathol. doi:10.1136/jclinpath-2012-200851
7. International Organization for Standardization. ISO 17043:2010 Conformity assessment—general requirements for proficiency testing. http://www.iso.org/iso/catalogue_detail.htm?csnumber=29366. Accessed 28 sept 2012
8. Thunnissen FB et al (2004) Dutch Pathology Molecular Diagnostic Working Group. Quality control in diagnostic molecular pathology in the Netherlands; proficiency testing for patient identification in tissue samples. J Clin Pathol 57:717–720
9. von Wasielewski R, Krusche CA, Rüschoff J, Fisseler-Eckhoff A, Kreipe H (2008) Implementation of external quality assurance trials for immunohistochemically determined breast cancer biomarkers in Germany. Breast Care 3:128–133
10. International Organization for Standardization. ISO 15189:2003 Medical laboratories—particular requirements for quality and competence. http://www.iso.org/iso/catalogue_detail?csnumber=42641. Accessed 28 Sept 2012
11. Rudiger T, Hofler H, Kreipe HH, Nizze H, Pfeifer U, Stein H, Dallenbach FE, Fischer HP, Mengel M, von Wasielewski R, Muller-Hermelink HK (2002) Quality assurance in immunohistochemistry: results of an interlaboratory trial involving 172 pathologists. Am J Surg Pathol 26:873–882
12. Mengel M, von Wasielewski R, Wiese B, Rüdiger T, Müller-Hermelink HK, Kreipe H (2002) Interlaboratory and interobserver reproducibility of immunohistochemical assessment of the Ki67-labelling index in a large multi-centre trial. J Pathol 198:292–299
13. Thunnissen E, Bovée JV, Bruinsma H, van den Brule AJ, Dinjens W, Heideman DA, Meulemans E, Nederlof P, van Noesel C, Prinsen CR, Scheidel K, van de Ven PM, de Weger R, Schuuring E, Ligtenberg M (2011) EGFR and KRAS quality assurance schemes in pathology: generating normative data for molecular predictive marker analysis in targeted therapy. J Clin Pathol 64:884–892
14. Andrew G, Carlin JB, Stern HS, Rubin DB (2003) Bayesian data analysis. Chapman & Hall, Boca Raton