

ORIGINAL ARTICLE

Quality control in diagnostic molecular pathology in the Netherlands; proficiency testing for patient identification in tissue samples

F B J M Thunnissen, M G J Tilanus, M J L Ligtenberg, P M Nederlof, W N M Dinjens, E Meulemans, A J C Van den Brule, C J M van Noesel, W J F de Leeuw, E Schuurin, on behalf of the Dutch Pathology Molecular Diagnostic Working Groups

J Clin Pathol 2004;57:717–720. doi: 10.1136/jcp.2003.011973

See end of article for authors' affiliations

Correspondence to:
Dr F Thunnissen, Canisius
Wilhelmina Hospital, Weg
door Jonkerbos 100, 6532
SZ Nijmegen, The
Netherlands;
e.thunnissen@
cwz.nl

Accepted for publication
2 January 2004

Aims: To describe the evolution of proficiency testing for molecular diagnostic pathology with respect to determining unambiguously the patient identity of tissue samples by microsatellite analysis.

Method: Four rounds of quality control exchanges of samples from different patients were sent with the purpose of identifying the correct origin of these samples. The samples were either paraffin wax embedded sections on glass, sections in tubes, or isolated DNA. Blinded samples were distributed to all participating laboratories. No restrictions to the method and short tandem repeat markers used for identification were imposed.

Results: In four subsequent rounds the number of participating laboratories increased from three to 10. The numbers of samples tested increased in time from five to 12. The microsatellite markers used by the different laboratories showed little overlap. In the first three rounds, in which isolated DNA was provided, all samples were accurately classified irrespective of the microsatellite markers used. In the last round, which also included paraffin wax embedded sections, a small number of laboratories experienced problems, either with amplification or incorrect classification of a few samples.

Conclusion: Proficiency testing was useful, and showed country wide high quality and correct identification of (patient) samples with molecular techniques for diagnostic purposes.

Quality control in laboratory testing has been performed in many fields in the past such as clinical pathology, microbiology, quantitative image analysis,^{1,2} flow cytometry, medical genetics,³ and recently molecular pathology.^{4,5}

The Dutch Pathology Society established a molecular diagnostic working group in 1998. A subcommittee organised external quality assurance (EQA) for the detection of sample identification by microsatellite analysis, the identification of p53 and K-ras mutations, polymerase chain reaction (PCR) based chromosomal translocations, IgH clonality assays, and human papillomavirus detection. In the past, proficiency testing in other areas has been performed several years after the initial development,^{1,2} leading to the realisation that essential basic improvements should be made. Therefore, we felt that proficiency testing in molecular laboratories would be useful in the implementation phase. Since 1998, five rounds of proficiency testing have been performed. The assays to be tested in the EQA were selected, based on information received from the pathology laboratories performing molecular diagnostic analysis in the Netherlands. In four of the five rounds of sample exchanges one of the subjects tested was the comparison of tissue samples from different patients.

“In the case of a major discrepancy between the endoscopic impression and the histopathological findings, short tandem repeat analysis may be performed to determine whether the biopsy is from the same patient”

In daily practice, this is sometimes done to determine the presence or absence of possible contamination or exchange of

samples.⁶ For example, in the case of a major discrepancy between the endoscopic impression and the histopathological findings, short tandem repeat (STR) analysis may be performed to determine whether the biopsy is from the same patient. The aim of our study was to report the outcome of proficiency testing in these four rounds of EQA for molecular diagnostics.

METHODS

Figure 1 shows the general scheme for one round of proficiency testing. A letter of invitation was sent from a central office to all pathology laboratories in the Netherlands. Each laboratory could participate on a voluntary basis. In each of the four rounds one of the subcommittee members organised the test samples. Split sample design was used to obtain several coded samples with equal composition from each specimen. Table 1 shows the composition of the test samples for the four rounds. In the case of paraffin wax embedded material, the samples were fixed in neutral buffered formaldehyde (4% wt/vol) according to the routine procedure. In the fourth round (2002), four samples (samples 3–6) were fixed for 24 hours and four samples were fixed in same fixative for six days (samples 7–10).

Coded test samples were sent to a distribution office for transfer to those who had reacted favourably to the letter of invitation. Two different people from the organising laboratory, who also participated in the proficiency testing, were involved in preparing and testing the samples. The first one prepared the coded sample and the second one was unaware of the coding and performed the testing.

Abbreviations: EQA, external quality assurance; PCR, polymerase chain reaction; STR, short tandem repeat

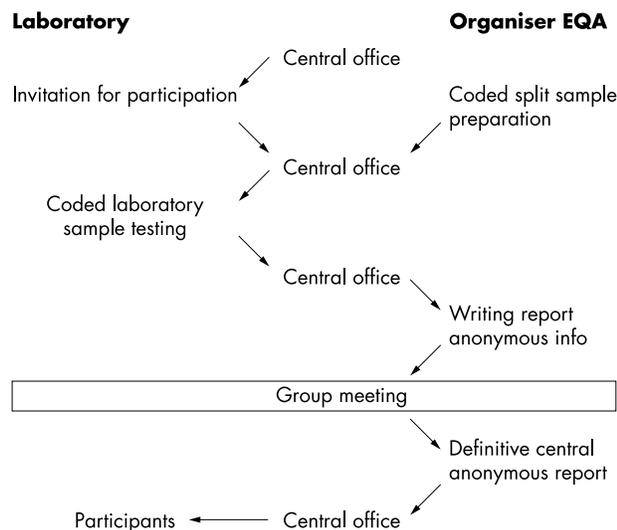


Figure 1 The general flow of activity for proficiency testing. The activities of the laboratories and the organiser EQA are listed under the appropriate headings. EQA, external quality assurance.

The results had to be returned within certain time limits accompanied by a questionnaire, containing the following questions:

- Which test sample(s) belong to the same patient(s)?
- Which markers were used, including a detailed technical protocol (standard operating procedure)?

Participating laboratories sent their results to the distributional office. The sample codes were disclosed and the organising subcommittee member compared the results from the coded samples with the original results (table 1). He was also responsible for summarising all the outcomes in a report.

The definitions of the terms used for the description of outcome were adapted and extended from the Association for Molecular Pathology, as follows⁷:

- Accuracy: ability to return a correct result compared with a reference.
- Precision: ability to produce the same result consistently on repeated testing of the same sample.
- Analytical specificity: ability to return a negative result when target nucleic acid is not present.
- Discordant: deviant from the reference.
- Compatible: not at variance with the reference, but not identical or not conclusive.

The compatible group contained cases where—for example, the extracted DNA could not be amplified; in other words, compatible determines for a sample the fraction of answers that still could be in agreement with the reference. In the case of 100% accuracy the use of the description compatible is not meaningful.

A concept report, with the laboratories still in coded form, was drawn up from the outcome, and was sent to all the participants. Each laboratory could establish its performance by comparison with the outcome regarded as the reference of each round. In addition, each laboratory could check at this point whether the report was correct with respect to its own data and compare its performance with the average or consensus. Subsequently, a central meeting was held, where the details of the outcome were discussed. Afterwards, a final report was made, with the anonymity of the laboratories being maintained. The procedure evolved during the quality

Table 1 The composition of the coded test samples for the four years

Sample	1998	1999	2000	2002
1	DNA A	DNA C	1 Par Bx E	DNA G
2	DNA B	0.001 DNA C	6 Par Bx E	DNA H
3	DNA B	DNA D	1 Par Rx E	1 Par H
4	Water	0.001 DNA D	6 Par Rx E	1 Par G
5	DNA A	Water	1 Par Rx F	6 Par G
6		Par C	DNA E	6 Par H
7		Par D		1 Par H
8				1 Par G
9				6 Par G
10				6 Par H
11				Glass G
12				Glass H

DNA denotes high molecular weight isolated DNA; 0.001 DNA indicates that the sample was diluted $\times 1000$ in Tris-EDTA. Par indicates that the tube contained one or two 3–4 μm paraffin wax embedded section(s) and Glass denotes a paraffin wax embedded section on a coated glass slide; each sample from the same patient contains the same letter (A–F). Bx and Rx indicate biopsy and resection, respectively. The numbers indicate the quantity compared with other samples in the same round. Samples 3–6 from 2002 were fixed for 24 hours and samples 9–10 for six days.

control rounds and the procedure described here was performed only in the last round. In the first rounds an additional reviewer went over the quality control report.

RESULTS

The number of participants increased during the four rounds from three, to six, eight, and 10. The markers used by the different laboratories varied greatly. The same markers were used in two laboratories only. Table 2 shows the markers used in 2002; in general, they are similar to those used in previous years. Table 1 shows details of the composition of the test samples in the four rounds.

In the first round the design of the coded test samples was simple. Only DNA from a few samples was distributed, providing information about the amplification and interpretation of electrophoresis from the STR markers only. All three laboratories obtained a correct classification of the samples. In the subsequent rounds the sample composition was slightly extended, with specific aims.

In the second round, two different concentrations of DNA and large paraffin wax sections in small tubes were sent out. This extension aimed to test DNA isolation from paraffin wax embedded sections and subsequent STR analysis. All six laboratories obtained a correct classification of the samples, irrespective of the DNA concentration. Extracting DNA from paraffin wax embedded sections in the different laboratories did not lead to a reduction in accuracy. In addition, in the first two rounds water samples were tested, which were classified as negative by PCR in all the laboratories, leading to a specificity of 100%.

In the third round, both one paraffin wax embedded section of a small biopsy and six sections of a paraffin wax embedded margin of a resection specimen from the same patient were sent out. Because all samples were classified accurately, it was clear that one section of a small biopsy was sufficient for STR analysis. This shows that these eight laboratories could correctly group tissue samples using one 10 μm thick paraffin wax embedded section from a 2–3 mm biopsy.

The fourth round contained the same elements as the previous rounds but, in addition, the samples had undergone variations in fixation time and DNA had to be extracted from a paraffin wax section that was mounted on a microscope slide. Table 3 shows the outcome of round four (2002). Six laboratories showed 100% precision when classifying samples

Table 2 The short tandem repeat markers used by different laboratories ranked by chromosome

Chromosome	Laboratory									
	1	2	3	4	5	6	7	8	9	10
1						D1S158				D1S158
2	D2S202				D2S1352		D2S1352	Tpox		D2S123
3		D3S1358		D3S1358	D3S2456		D3S2456			D3S1298
4	D4S174	FGA		FGA		D4S430				
5	D5S346	D5S818		D5S818		D5S346		CSF		D5S346
6							D6S624			
7		D7S820		D7S820						
8		D8S1179		D8S1179			D8S858		D8S133, S255, S261	
9			D9S171, S156		D9S932	D9S63	D9S932		D9S156, S171, S176	D963
11	D11S875							TH0	hAR	
12		VWA		VWA				VWA		
13		D13S317	D13S1304, S153	D13S317						
14			D14S977			D16S26224		D16S2624		
17	D17S250									D17S250
18		D18S51		D18S51	D18S858	D18S58		D18S51		D18S58, S61
19					D19S565		D19S565			
21		D21S11		D21S11				D21S11		
X,Y		Amelogenin		Amelogenin						

Additional information including primers can be obtained from web sites (for example, <http://www.gdb.org> and <http://www.ncbi.nlm.nih.gov>).

from the same patient. However, two of the 10 laboratories had technical problems with amplifying DNA for seven of 120 scores (6%). This outcome is essentially not discordant and was scored as compatible because the conclusion based upon the results obtained was not incorrect. The problems with PCR could not be attributed to prolonged fixation time, because the accuracy of classification was better for samples with longer fixation. In one laboratory, four samples were classified as discordant. In another laboratory, one of the samples, where DNA had to be extracted from the paraffin wax embedded section on a glass slide, was not conclusive and was incorrectly classified. In total, five of the 120 scores were incorrectly classified (4%). Overall, the score for this round was less accurate (96% compatible) than in the first three rounds (100%).

DISCUSSION

The results of proficiency testing were excellent in the first three rounds. A highly accurate classification was achieved, irrespective of the choice and number of markers used. In the first three rounds all samples were classified with 100% accuracy. It was clear that for a standard procedure of testing at least one section was sufficient for unambiguous sample

identification. Thus, sample exchange identification using STR analysis does not require extensive molecular equipment and can be done by most laboratories that undertake molecular diagnostic analyses. In the 2002 round of proficiency testing (the fourth round), two of the coded test samples were paraffin wax embedded sections on glass slides, therefore extending the analysis to include DNA extraction from glass slides. From the discussion meeting after the concept report, it appeared that more laboratories had difficulties with PCR, but most were able to classify these samples correctly. However, one laboratory was not able to resolve the PCR or DNA extraction related problems and another was not sure why one sample was classified incorrectly. Because this laboratory accurately classified all the other samples, the specific improvement needed in this laboratory is probably the isolation of DNA from glass slides. Remarkably, in one laboratory four samples were classified as discordant. This was unexpected and might have resulted from the samples being mixed up early in the process of DNA isolation at the test site. Although samples were prepared in the distributing laboratory in a manner aimed at minimising mistakes and mix ups, we cannot exclude this as a reason for the outcome in this single laboratory. Because of these results

Table 3 Outcome of fourth round of proficiency testing

Sample	Reference	Laboratory									
		1	2	3	4	5	6	7	8	9	10
1	A DNA	A	A	A	A	A	A	A	A	A	A
2	B DNA	B	B	B	B	B	B	B	B	B	B
3	B Par	B	B	B	B	B	B	B	B	B	B
4	A Par	A	A	NP	A	A	A	A	A	A	A
5	A Par	A	A	NP	A	A	NP	A	A	A	A
6	B Par	B	B	NP	B	B	B	B	B	B	A
7	B Par	B	B	B	B	B	B	B	B	B	A
8	A Par	A	A	A	A	A	A	A	A	A	B
9	A Par	A	A	NP	A	A	A	A	A	A	B
10	B Par	B	B	B	B	B	B	B	B	B	B
11	A Glass	A	A	NP	A	A	A	A	A	B?	A
12	B Glass	B	B	NP	B	B	B	B	B	B	B

DNA denotes high molecular weight isolated DNA. Par indicates that the tube contained one or two 3–4 µm paraffin wax embedded section(s) and Glass denotes a paraffin wax embedded section on a coated glass slide; each sample from the same patient contains the same letter (A or B). NP, no polymerase chain reaction product; ?, uncertain classification. Bold indicates discordant results.

in future we will store some spare samples for laboratories that might want to re-evaluate their testing. The outcome of the fourth round also shows that proficiency testing is useful even for seemingly simple tests. Overall, the results for patient identification were good, particularly because these techniques have only recently been developed and are rapidly applied. This may in part be explained by the fact that although many markers are used, only one or a few are necessary for adequate distinction.

Most laboratories used five or six markers for STR analysis. In general, this was sufficient for correct classification in all four rounds, because relevant differences between the samples were obtained. In a previous publication, tissue identification was reported with radioactive labelling of X and Y chromosome probes.⁶ In our report identification was performed with ethidium bromide staining in agarose gels and later with increasing frequency by fluorescence labelling, as a result of technical improvements with time.

During the central meeting at the end of each round of proficiency testing the laboratory coding was voluntarily decoded by the participants, and the data were discussed with respect to the outcomes from the different laboratories. This resulted in the exchange of protocols where appropriate, such as DNA isolation from paraffin wax embedded material.

The time frame of the tests was frequently a topic of discussion at the evaluation meeting. Although in daily practice most laboratories could provide an answer for STR analysis (including DNA analysis) in 4.7 days (range, 3–7), several laboratories requested that the testing time should be extended further than the initial four to six weeks. This was partly because of the low priority given to this testing in the work of the laboratory. Therefore, the importance of quality control does not yet seem to be firmly established in routine molecular diagnostics.

Performing this type of quality control early in the implementation phase of new molecular techniques for pathology showed that this technique is adequately performed in many laboratories, even if the initial experience is limited. This is a promising feature of molecular pathology testing, in contrast to some other techniques.^{1,2}

In the Netherlands there is no legal obligation to participate in proficiency testing. The Dutch Pathology Society encourages laboratories to participate in molecular proficiency testing. No sanctions exist for laboratories performing below the average laboratory performance. However, it is felt that such laboratories would be eager to show improved performance in the subsequent rounds of proficiency testing. This lack of enforcement and the collaborative stimulating attitude has led to the situation where all laboratories performing molecular pathological diagnostic tests participate in EQA on a voluntary basis. However, because most diagnostic assays should fulfil minimal standards, quality control is essential. Therefore we strongly encourage laboratories to join because we foresee that in coming years national EQA testing will become obligatory. In general, to improve laboratory testing a laboratory needs to perform and compare its local assays with those of others. Participating in EQA and performing at an optimal level proves that the laboratory is fulfilling its responsibility. The participants set the standard and the peer review activity of the network members takes care of optimal standards.

In conclusion, the proficiency testing of STR analysis for possible exchange of tissue samples shows that molecular

Take home messages

- Proficiency testing of short tandem repeat analysis for possible exchange of tissue samples shows that molecular techniques are performed accurately in many laboratories
- This outcome is independent of the markers used
- The central meeting afterwards enables standardisation of protocols, where appropriate
- More laboratories had difficulties with DNA extraction from glass slides when using other sources of DNA

techniques can be performed in many laboratories with high accuracy. This outcome is independent of the markers used. The central meeting afterwards enables standardisation of protocols, where appropriate.

ACKNOWLEDGEMENTS

The help of “Stichting Kwaliteitscontrole Klinische Pathologie” for the distribution of the samples in the last round is greatly appreciated.

Authors' affiliations

F B J M Thunnissen, Department of Pathology, Canisius Wilhelmina Hospital, 6532 SZ Nijmegen, The Netherlands

M G J Tilanus, Department of Pathology, University Medical Centre, 3508 GA Utrecht, The Netherlands

M J L Ligtenberg, Department of Pathology, University Medical Centre, 6500 HB Nijmegen, The Netherlands

P M Nederlof, Department of Pathology, Antoni van Leeuwenhoek Hospital, 1066 CX Amsterdam, The Netherlands

W N M Dinjens, Department of Pathology, Erasmus University Medical Centre, 3000 DR Rotterdam, The Netherlands

E Meulemans, Department of Pathology, Maastricht University Medical Centre, 6202 AZ Maastricht, The Netherlands

A J C Van den Brule, Department of Pathology, Free University Medical Centre, 1007 MB Amsterdam, The Netherlands

C J M van Noesel, Department of Pathology, Academic Medical Centre, 1105 AZ Amsterdam, The Netherlands

W J F de Leeuw, Department of Pathology, Leiden University Medical Centre, 2300 RC Leiden, The Netherlands

E Schuurin, Department of Pathology, Groningen University Medical Centre, 9713 EZ Groningen, The Netherlands

REFERENCES

- 1 **Thunnissen FB**, Ellis IO, Jutting U. Quality assurance in DNA image analysis on diploid cells. *Cytometry* 1997;**27**:21–5.
- 2 **Thunnissen FB**, Ellis IO, Jutting U. Interlaboratory comparison of DNA image analysis. *Anal Cell Pathol* 1996;**12**:13–24.
- 3 **Muller CR**. Quality control in mutation analysis: the European Molecular Genetics Quality Network (EMQN). *Eur J Pediatr* 2001;**160**:464–7.
- 4 **Hsi ED**, Tubbs RR, Lovell MA, et al. Detection of bcl-2/1(H) translocation by polymerase chain reaction: a summary of the experience of the molecular oncology survey of the College of American Pathologists. *Arch Pathol Lab Med* 2002;**126**:902–8.
- 5 **Cabras AD**, Kremer M, Schulz S, et al. Quality assessment in diagnostic molecular pathology: experience from a German–Austrian–Swiss multicenter trial. *Virchows Arch* 2000;**437**:46–51.
- 6 **Worsham MJ**, Wolman SR, Zarbo RJ. Molecular approaches to identification of tissue contamination in surgical pathology sections. *J Mol Diagn* 2001;**3**:11–15.
- 7 **Association for Molecular Pathology statement**. Recommendations for in-house development and operation of molecular diagnostic tests. *Am J Clin Pathol* 1999;**111**:449–63.