

ARTICLE

Improvement of interpretation in cystic fibrosis clinical laboratory reports: longitudinal analysis of external quality assessment data

Sarah Berwouts^{1,6}, Emmanuelle Girodon^{2,6}, Martin Schwarz^{3,6}, Manfred Stuhmann^{4,6}, Michael A Morris^{5,6} and Elisabeth Dequeker^{*,1,6}

Participation in external quality assessment (EQA) is a key element of quality assurance in medical laboratories. In genetics EQA, both genotyping and interpretation are assessed. We aimed to analyse changes in the completeness of interpretation in clinical laboratory reports of the European cystic fibrosis EQA scheme and to investigate the effect of the number of previous participations, laboratory accreditation/certification status, setting and test volume. We distributed similar versions of mock clinical cases to eliminate the influence of the difficulty of the clinical question on interpretation performance: a cystic fibrosis patient (case 1) and a cystic fibrosis carrier (case 2). We then performed a retrospective longitudinal study of reports over a 6-year period from 298 participants for case 1 (2004, 2008, 2009) and from 263 participants for case 2 (2006, 2008, 2009). The number of previous participations had a positive effect on the interpretation score ($P < 0.0001$), whereas the laboratory accreditation/certification status, setting and test volume had no effect. Completeness of interpretation improved over time. The presence of the interpretation element 'requirement for studying the parents to qualify the genotype' increased most (from 49% in 2004 to 93% in 2009). We still observed room for improvement for elements that concerned offering testing for familial mutations in relatives and prenatal/preimplantation diagnosis (16% and 24% omission, respectively, for case 1 in 2009). Overall, regular participation in external quality assessment contributes to improved interpretation in reports, with potential value for quality of care for patients and families by healthcare professionals involved in genetic testing.

European Journal of Human Genetics (2012) 20, 1209–1215; doi:10.1038/ejhg.2012.131; published online 20 June 2012

Keywords: clinical interpretation; quality improvement; cystic fibrosis; external quality assessment; laboratory report

INTRODUCTION

Constitutional genetic information is constant throughout life; patients are typically tested once in their lifetime and the test may never be repeated or confirmed. Consequently, incorrect results may not become apparent for years and can lead to harmful outcomes for patients and family members, such as unnecessary surgery or prenatal diagnosis, or erroneous carrier testing. Given the complex and sensitive nature of genetic information, testing services have the responsibility to accurately test and interpret genetic results in the context of the clinical indication and family history, while taking into account that the recipient of the laboratory report may not have genetic expertise.

In the 1990s, a system known as external quality assessment (EQA, in Europe) or proficiency testing (PT, in the United States) was introduced in molecular genetic diagnostic laboratories to evaluate the ability to accurately determine genotypes and to assess the accuracy and completeness of the interpretation of the test results.¹ Independent assessment of test performance is a recommendation of the Organisation for Economic Cooperation and Development, and a requirement for laboratory accreditation to the International

Organisation for Standardisation (ISO) 15189 and 17025 standards and for complying with the US Clinical Laboratory Improvement Amendments (CLIA).^{2–5} One of the main European EQA providers in molecular genetics is the Cystic Fibrosis Network, which has been organising an EQA scheme for cystic fibrosis (CF; OMIM no. 219700) since 1996.⁶

Assessment of reports through the Cystic Fibrosis Network and other disease-specific schemes organised by the European Molecular Genetics Quality Network revealed that there is room for improvement of the interpretation in laboratory reports. For example, the EQA scheme for molecular diagnosis of hereditary recurrent fevers described that only 8–47% of the reports mentioned interpretation of testing results (15–50 participants, 2006–2008).⁷ Similarly, the spinocerebellar ataxias EQA scheme reported average interpretation scores of 1.53/2.00 and 1.76/2.00 (26–37 participants, 2005–2006) and the scheme for mutation detection in the breast cancer genes revealed average interpretation scores from 1.46/2.00 to 1.78/2.00 (25–41 participants, 2000–2002).^{8,9} The maximum interpretation score in these schemes was 2.0 per laboratory; the average interpretation scores represented the average of interpretation scores of all participating laboratories in a certain year.

¹Department of Public Health, Biomedical Quality Assurance Research Unit, Katholieke Universiteit Leuven, Leuven, Belgium; ²Service de Biochimie et Génétique, Groupe hospitalier Henri Mondor, Créteil, France; ³Genetic Medicine, St Mary's Hospital, Manchester, UK; ⁴Institut für Humangenetik, Medizinische Hochschule, Hannover, Germany; ⁵Molecular Diagnostic Laboratory, Service of Genetic Medicine, University Hospitals, Geneva, Switzerland

*Correspondence: Professor E Dequeker, Department of Public Health, Biomedical Quality Assurance Research Unit, Katholieke Universiteit Leuven, Herestraat 49, box 602, 3000 Leuven, Belgium. Tel: +32 16 34 58 81; Fax: +32 16 34 79 49; E-mail: els.dequeker@med.kuleuven.be

⁶All the authors are members of the Cystic Fibrosis European Network.

Received 16 January 2012; revised 22 May 2012; accepted 22 May 2012; published online 20 June 2012

Although EQA is accepted as a corner stone of quality assurance, direct evidence that participation in EQA leads to improved interpretation is lacking because measuring improvement of interpretation is challenging.^{10,11} First, interpretation depends heavily on the complexity of the mock clinical cases, which change every year.¹² Second, published interpretation scores represent the mean annual score for the whole group of participants, whose members vary from year to year, with a mix of regular, irregular and new participants.

This large retrospective longitudinal study was performed through the Cystic Fibrosis Network over a 6-year period among more than 300 laboratories. We examined and analysed changes in the completeness of interpretation in clinical laboratory reports when similar versions of mock clinical cases were distributed and thus when the influence of the complexity of the clinical question was eliminated. In addition, we investigated the effect of the number of previous EQA participations on the completeness of the interpretation and thus evaluated the efficacy of EQA. Monitoring and improving the completeness of interpretation in clinical reports is essential as it will lead to better quality of care for patients and families.

MATERIALS AND METHODS

Context and setting of the study

In a standard CF EQA round (which takes 1 year, starting from autumn until summer the year after), a set of samples (purified DNA) and mock cases for closely related clinical situations are sent for analysis to voluntarily registered laboratories. Laboratory reports submitted by the participants to the EQA provider should reflect routine reports issued to requesting physicians. The assessors determine the key interpretation elements that should be present in participants' laboratory reports before dispatching the samples. Their decisions are based on expert consensus and best practice guidelines, and depend on the clinical indication in the case.^{13–16} Two assessors independently evaluate genotypes and interpretation in the submitted reports. Results are also discussed during an assessment meeting. Apart from laboratory reports, the CF Network requests information on the accreditation and certification status of the laboratory, the laboratory setting (eg, hospital, university, industry) and the number of CF samples tested per year (laboratory test volume). At the end of the EQA round, participants receive feedback in the form of laboratory-specific individual comments and a general report addressed to all participants. Laboratory information and performance data have been stored systematically in a database since 2004, using a unique identification code for each laboratory.

Study design

Over a 6-year period, we have intentionally included similar versions of two mock cases, presenting clinical situations requiring essentially the same interpretation, masked by changes to the case descriptions (patients' names, genders, ethnic origins and dates of birth) and the genotypes: three versions of a patient with CF in 2004, 2008 and 2009, named case 1, and three versions of a CF carrier in 2006, 2008 and 2009, named case 2 (Table 1). The genotypes were chosen not to affect the evaluation. The laboratories were not informed about the similarity of the cases, nor that we were conducting this study.

The interpretation elements evaluated in the clinical laboratory reports for both cases in the study are outlined in Table 2. A laboratory received marks when an interpretation element was present and correct in the report, or when an element was present but not clearly defined (eg, a laboratory recommends genetic counselling without specifying whom it may concern). No marks were awarded when the element was absent or if the element was present, but wrong (eg, a wrong risk figure). The sum of the marks (further referred to as the 'interpretation score') is maximum 2.0 per case.

We collected the stored performance data for these specific similar cases only at the end of the 6-year period and then conducted a longitudinal retrospective analysis. If participants provided laboratory information (accreditation/certification status, setting and test volume) in different years, only the most recently submitted information was used in this study.

Laboratories that registered for the EQA scheme but did not submit laboratory reports were not taken into account in the study (eight laboratories in 2004, five in 2006, four in 2008 and six in 2009). We also excluded those reports in which a genotyping error was made, because interpretation was not assessed the same way (one laboratory in 2004, four in 2008, five in 2009 for case 1, and six in 2006, three in 2008, seven in 2009 for case 2). Furthermore, we excluded reports from laboratories that did not report a certain mutation because they had not tested for it (two laboratories in 2006, three in 2008, eight in 2009 for case 1, and one in 2008, one in 2009 for case 2).

Statistical analysis

The interpretation score for a laboratory comprised values between 0.0 and 2.0 with nine possible levels. Given the very skew distribution of the score (higher scores are more frequent), we adopted a proportional odds model (odds ratios (OR) and 95% confidence intervals (CI)) to analyse the effect of the number of previous EQA participations, the effect of the year, the laboratory accreditation/certification status, setting and test volume. The proportional odds model considered the interpretation score as an ordered category response and modelled the probability that the response fell within the higher categories (equivalent to a higher score) *versus* the lower categories (lower score). A *P*-value <0.05 was considered significant. Generalised estimating equations were used for estimation to account for clustering of observations coming from the same laboratory.

The specific interpretation elements on which the laboratory reports were evaluated were given a binary score (0/1; element absent or element wrong/element present and correct or present but not clearly defined). The effect of the number of previous participations on the success probability on each aspect were analysed using a logistic regression model and using generalised estimating equations to take into account the clustering of the scores due to the longitudinal data structure.

All analyses have been performed using SAS software, version 9.2 of the SAS System for Windows (SAS Institute Inc., Cary, NC, USA).

RESULTS

Study population

Data regarding the characteristics of the 311 laboratories (272 European), from 39 different countries included in the study are described in Table 3.

All these laboratories offered testing for cystic fibrosis transmembrane conductance regulator (*CFTR*) gene mutations as a clinical test. We assessed the interpretation in clinical reports from 298 different laboratories for case 1 ($n=232$ in 2004, $n=202$ in 2008, $n=203$ in 2009) and from 263 different laboratories for case 2 ($n=200$ in 2006, $n=205$ in 2008, $n=208$ in 2009). Roughly half of the laboratories provided reports that could be analysed in each of the 3 investigated years: 44% (132/298) for case 1 and 56% (147/263) for case 2. About one-third of the laboratories provided reports in only one of the three investigated years: 31% (91/298) for case 1 and 23% (60/263) for case 2.

Trend of the mean annual interpretation score

First, we calculated the mean annual interpretation score for the whole group of participants for both cases (black rounds in Figure 1). In general, scores improved over the years. A significant difference was found between the two cases with better interpretation scores for case 1 ($P=0.0007$; OR: 1.41; CI: 1.16–1.72). For case 1 in 2004, only 20% (46/232) of the laboratories got a maximum interpretation score of 2.0, compared with 66% (134/203) in 2009. A similar trend was observed for case 2: 31% (61/200) got a maximum score of 2.0 in 2006 and 61% (126/208) in 2009. We then analysed whether the laboratories that participated more frequently in the CF EQA scheme had better scores than those that participated less frequently. The score for the laboratories that participated three times or only once

Table 1 Overview of the similar versions of two cases, sent to the laboratories that participated in the CF EQA scheme in 2004 or 2006, in 2008 and in 2009

Year	Identification	Clinical indication	Genotype (HGVS/traditional)
<i>Case 1</i>			
2004	Emily Thibault Brussels Belgium 15 August 2004	Emily Thibault is a neonate who had a meconium ileus at birth. The paediatrician suspected CF and requests <i>CFTR</i> testing for confirmation of diagnosis. She is of Belgian origin and has no family history of CF.	c.[1521_1523delCT-T];[1521_1523delCTT] (F508del/F508del)
2008	Pieter van Hulle Nijmegen The Netherlands 31 July 2008	Pieter had meconium ileus at birth. There is no history of CF on either side of the family. A molecular analysis is requested for confirmation of the diagnosis of CF. Pieter is of Dutch origin.	c.[1652G>A];[1657C>T] (G551D/R553X)
2009	Fritz Ackermann Leipzig Germany 10 September 2009	Fritz is a newborn who had meconium ileus at birth. The paediatrician suspects CF and requests a <i>CFTR</i> gene molecular analysis to confirm the diagnosis. There is no family history of CF.	c.[489 + 1 G>T];[1521_1523delCTT] (621 + 1G>T/F508del)
<i>Case 2</i>			
2006	Carol Robertson Manchester United Kingdom 01 February 1979	Carol Robertson had a sister who died of CF in early infancy. She is planning a pregnancy and would like to know her carrier status and has specifically requested her risk of having a CF child. No molecular study has previously been done in the family. She is from the UK and her partner is from Spain. He is in good health and has no known family history of CF.	c.[1519_1521delATC];[=] (1507del/normal)
2008	William Lewis Glasgow United Kingdom 08 June 1977	Florence and her husband William are planning their first pregnancy. William is in good health, but had a younger sister who died of CF some years ago and he would like to know his carrier status and the risk of having a CF child. No molecular studies have been done in the family. Florence has no family history of CF and is in good health. She is from France, William is from Scotland.	c.[1585-1G>A];[=] (1717-1G>A/normal)
2009	Jack Hurst Birmingham United Kingdom 08 November 1983	Jack, a healthy 26-year-old man, has a sister who died of CF in early infancy. He and his partner are planning a pregnancy and he would like to know his carrier status and his risk of having a CF child. No molecular study has previously been done in the family. His partner, of British origin, is in good health and has no history of CF.	c.[1519_1521delATC];[=] (1507del/normal)

Table 2 Overview of the interpretation elements evaluated in the clinical laboratory reports for both cases in the study

Interpretation element	Mark
<i>Case 1</i>	
Clear confirmation of the diagnosis of cystic fibrosis	1.00
Mention qualification of the genotype is needed (= mention that both parents need to be tested to confirm compound heterozygosity)	0.25
Suggest or offer genetic counselling for the couple	0.25
Suggest or offer prenatal diagnosis or preimplantation genetic diagnosis for the next pregnancy	0.25
Suggest or offer cascade screening in relatives (= mention that testing for the familial mutation(s) should be offered to relatives)	0.25
<i>Case 2</i>	
Clear confirmation of the carrier status of cystic fibrosis	0.50
Suggest or offer to test the partner of the cystic fibrosis carrier	0.50
Suggest or offer genetic counselling for the couple	0.25
Risk figure of having a child with cystic fibrosis	0.50
Suggest or offer cascade screening in relatives	0.25

during the investigated period are indicated with grey symbols in Figure 1. We observed a significant positive effect of the number of previous participations on the interpretation score, both for case 1

($P < 0.0001$; OR: 2.34; CI: 1.60–3.44) and for case 2 ($P < 0.0001$; OR: 2.64; CI: 1.85–3.76).

We analysed the effect of the year on the interpretation score and corrected for the number of previous participations. A lower score was observed for 2004 compared with 2008 in case 1 ($P < 0.0001$; OR: 0.43; CI: 0.28–0.65), and a lower score was observed in 2009 compared with 2008 in case 2 ($P = 0.017$; OR: 1.67; CI: 1.10–2.54). No effect of the year was observed between 2008 and 2009 in case 1, and between 2006 and 2008 in case 2.

Trend of the presence of specific interpretation elements

We visualised the trend of the presence or absence of the specific interpretation elements for both cases between 2004 and 2009 (Figure 2). The presence of all elements increased steadily over the years, except 'confirmation of CF carrier status', which decreased in 2009 (86%, 178/208) compared with 2008 (97%, 199/205) and 'genetic counselling for the couple' for case 2, which decreased slightly in 2008 (85%, 175/205) compared with 2006 (88%, 175/200). Note that we accepted both 'elements present' and 'elements present, but not clearly identified' as 'present' in this analysis. The presence of interpretation element 'need for qualification of the genotype' (that is, study the parents to confirm homozygosity or compound heterozygosity) increased most, from 49% (113/232) in 2004 to 93% (189/203) in 2009. 'Cascade screening for relatives'

increased also more than 30% (from 48% (111/232) to 84% (171/203) in case 1 and from 46% (91/200) to 78% (162/208) in case 2). At the end of the study, two elements were still relatively absent in the clinical reports: the element 'cascade screening in relatives', as mentioned just above, and the element 'offer or suggest prenatal diagnosis or preimplantation genetic diagnosis for next pregnancy', which was omitted from 24% (49/203) of the reports in 2009 (46% (107/232) in 2004). As with the interpretation scores,

Table 3 Characteristics of 311 laboratories, from 39 countries included in the study

Characteristic	Case 1 n (%)	Case 2 n (%)
Participation		
Countries ^a	39	37
Laboratories	298	263
Accreditation/certification status of the laboratory^b		
Accredited/certified before 2004	38 (13%)	39 (15%)
Accredited/certified between 2004 and 2007	44 (15%)	52 (20%)
Accredited/ certified since 2008	47 (16%)	48 (18%)
No accreditation/no certification	124 (41%)	124 (47%)
Unknown	45 (15%)	—
Laboratory setting		
Hospital/university/university hospital	184 (62%)	157 (60%)
Private/industry	93 (31%)	91 (35%)
Other	16 (5%)	11 (4%)
Unknown	5 (2%)	4 (1%)
Number of CF samples tested per year (laboratory test volume)		
1–99	95 (32%)	88 (33%)
100–249	63 (21%)	63 (24%)
250–999	83 (28%)	70 (27%)
> 1000	49 (16%)	42 (16%)
Unknown	8 (3%)	—

^aAustralia, Austria, Belgium, Bulgaria (not for case 2), Canada (not for case 2), Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Israel, Italy, Latvia, Lithuania, Macedonia, New-Zealand, Norway, Poland, Portugal, Qatar, Russia, Saudi-Arabia, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, The Netherlands, Turkey, Ukraine, United Kingdom and the United States.

^bWe categorised for this study the following standards and programs under accreditation: ISO 15189, ISO 17025, Clinical Pathology Accreditation, Coördinatie Commissie ter bevordering van de Kwaliteitsbeheersing op het gebied van Laboratoriumonderzoek in de Gezondheidszorg, College of American Pathologists accreditation, Clinical Laboratory Improvement Amendments and Joint Commission International. Certification standard: ISO 9001.

we observed a positive effect of the number of previous participations on the presence of all specific interpretation elements, except for 'confirmation of the CF carrier status' ($P=0.47$) in case 2 (Table 4).

Effect of the laboratory accreditation or certification status, setting and test volume on the interpretation

We evaluated whether laboratories that were accredited or certified had more complete interpretation than those that were not accredited or certified. In addition, we evaluated whether a certain setting was associated with better interpretation and whether laboratories that test more CF samples per year had better scores than those that test fewer samples per year. No effect was observed of the laboratories' accreditation/certification status on the interpretation scores ($P=0.18$ for case 1, $P=0.27$ for case 2). However, within the group of accredited and certified laboratories, there was an effect of the number of years since the accreditation or certification. We found lower scores for laboratories that had been accredited or certified for a longer time for case 1 ($P=0.0004$; OR: 0.90; CI: 0.85–0.95) and case 2 ($P=0.005$; OR: 0.93; CI: 0.89–0.98). For example, the 2009 mean annual interpretation score for case 1 was 1.77/2.00 for the laboratories that had achieved accreditation before 2004, 1.83/2.00 for the laboratories accredited between 2004 and 2007 and 1.90/2.00 for the laboratories accredited since 2008. The outcome was similar when the analysis was restricted to accredited laboratories. No differences in interpretation score were found between the three different categories of laboratory settings ($P=0.99$ for case 1, $P=0.63$ for case 2) and between the four different categories of laboratory test volumes ($P=0.71$ for case 1, $P=0.2$ for case 2).

DISCUSSION

Principal findings

First, this study with similar mock clinical cases revealed that regular participation in EQA leads to a more complete interpretation in laboratory reports. Our observations support the view that EQA participation is a form of continuous education, whereby a comprehensive EQA general report and individual comments contribute to the improvement observed. Interpretation could additionally be influenced by external factors such as the availability of best practice guidelines and publications.^{13–15}

Second, the presence in clinical reports of all investigated interpretation elements improved over the years. Two elements increased remarkably more (>30%) than others: stating the need to test the parents of a CF patient to confirm homozygosity or compound heterozygosity (need for qualification of the genotype) and

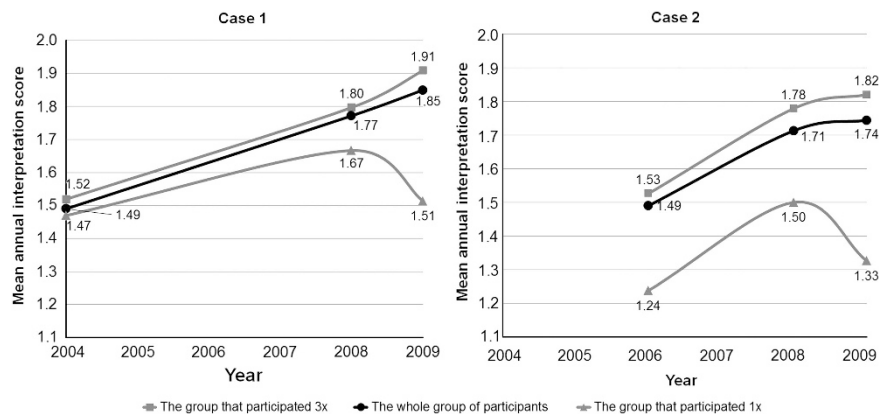


Figure 1 Trend of the interpretation scores for different groups for case 1 (2004, 2008 and 2009) and case 2 (2006, 2008 and 2009).

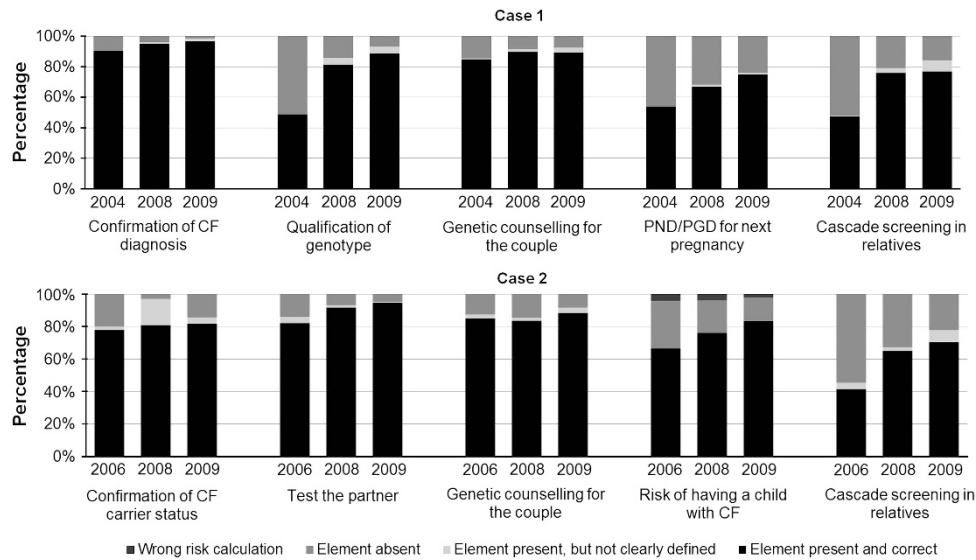


Figure 2 Presence or absence of specific interpretation elements, in the laboratory reports for case 1 (2004, 2008 and 2009) and case 2 (2006, 2008 and 2009).

Table 4 Effect of the number of previous EQA participations on the presence of specific interpretation elements

Interpretation elements	P-value	OR	95% CI
Case 1			
Confirmation of CF diagnosis	0.0266	3.08	1.14–8.30
Qualification of genotype	0.0002	2.85	1.63–4.97
Genetic counselling for the couple	0.0071	2.48	1.28–4.81
PND/PGD for next pregnancy	<0.0001	2.23	1.57–3.17
Cascade screening in relatives	0.0423	1.65	1.02–2.69
Case 2			
Confirmation of CF carrier status	0.4651	—	—
Test the partner	0.0003	3.15	1.69–5.85
Genetic counselling for the couple	<0.0001	3.92	2.20–6.99
Risk figure of having a child with CF	<0.0001	3.31	2.19–5.00
Cascade screening in relatives	0.0003	2.04	1.38–3.01

Abbreviations: CI, confidence interval; OR, odds ratio. P-values < 0.05 were considered significant.

mentioning the possibility of testing the relatives. The fact that most laboratories currently mention that genotypes must be confirmed represents an important improvement in laboratory practice and a reduced probability of error or misinterpretation, and shows that laboratories better understand the limitations of genetic analysis, especially PCR-based tests.

Third, although all investigated interpretation elements were included in the majority (>80%) of reports by the end of the study, there was room for improvement for those aspects that concern offering testing for familial mutations in relatives or future pregnancies. The possibility of testing relatives should be mentioned in all cases in which an individual is positive for a CF-causing mutation. It is also recommended that laboratories mention the possibility of offering prenatal diagnosis for future pregnancies in all cases where a child was born with CF, as the risk of a future child with CF is one-fourth in each pregnancy. Opportunities for improvement were noticed more for case 2 than for case 1. This may be because

laboratories interpret test results of patients with CF more assiduously than results of CF carriers. Nevertheless, in either case, genetic test reports can be referred to at a later date for relatives and future generations, and both situations must therefore be treated with the same assiduity.

Finally, no effect was observed of the laboratory accreditation/certification status on the completeness of the interpretation. Laboratories that are not accredited, but that perform EQA, reach an equal quality of their clinical reports compared with accredited or certified laboratories, confirming EQA participation as a fundamental element of laboratory quality assurance. The reason why recently accredited/certified (accredited) laboratories performed significantly better compared with those with more years of accreditation/certification (accreditation) is unclear. There might be a chance that recently accredited laboratories pay more attention to the improvement of their reports as they still have regular audits, whereas laboratories that have their quality management system in place for a long time, may become complacent. They perhaps just maintain their systems, with less attention to major adaptations.

Strengths and limitations

Our study has several important strengths. First, this is the largest longitudinal expert peer review of interpretation in genetic reporting for a single disorder. We were able to take advantage of the unique opportunity of having a large number of laboratories involved, from Europe, the United States and Australasia, principally because CF is a very commonly tested disorder compared with other genetic diseases. To evaluate the proportion of *CFTR*-testing laboratories in this study, we consulted Orphanet and estimated that 75% (272/361) of listed *CFTR*-testing laboratories were included in our study.¹⁷ Second, this is the first study that intended to overcome the influence of the difficulty of the clinical case on the interpretation score by including similar cases over a relatively long period. Further, difference in interpretation scores were analysed statistically. Published results of other EQA schemes have mainly included data from EQA pilot years, without any statistical evidence that the interpretation improved.^{7–9} It is known that assessment criteria and procedures are being set up in the pilot years and this could have a considerable impact on the trend

of the interpretation score. In comparison, the CF EQA scheme was set up 8 years before this study and with a fixed group of assessors over the years.

Although this study provides important data on the impact of the number of previous participations on the interpretation in laboratory reports, it has potential limitations. Firstly, only reports for the EQA scheme were evaluated, and it is not possible to judge whether the laboratories' real practice has improved, or if tailored reports are prepared to 'satisfy' the EQA provider'. Secondly, although we have eliminated the influence of the variety of the clinical question, there might be a variety in the strictness of assessment. Even though the assessors group did not change over the years, it is likely that assessment was stricter year after year, thus reducing interpretation scores at later years. There might be a chance that the group of laboratories that participated only one time during the investigated period experienced a greater effect of the increase of strictness, compared with those that participated more times (cf. decrease of the mean annual interpretation score in Figure 1 for the group that participated one time). Further, the data on laboratory accreditation/certification status were provided by the laboratories themselves. We know from experience within the European project EuroGentest that misunderstanding about the exact meaning of accreditation is common and we cannot rule out that some laboratories claimed to be accredited, but meant licensed, or mistook hospital accreditation for laboratory accreditation and thus overestimated their status.^{18,19} The outcome with regard to the accreditation and certification status could have been different if the information given by the laboratories was 100% correct. Seen the fact that the accurateness of the data could not be guaranteed, we decided not to perform detailed research on the accreditation/certification status data. We thus did not analyse for instance whether or not the effect of accreditation/certification status on the interpretation score is generalisable to all accreditation and certification programs.

CONCLUSION AND IMPLICATIONS

A genetic test result differs from other diagnostic results in that it can have far-reaching effects on patients' and their relatives' life plans; furthermore, genetic tests are usually performed only once in a lifetime. Consequently it is of utmost importance that the result is correct and accompanied with an unambiguous, accurate and complete interpretation. We demonstrate that regular participation in EQA contributes to a more complete interpretation in laboratory test reports. Unfortunately, 50% of the CF EQA participants still do not participate regularly in the CF EQA scheme. In addition, based on data from Orphanet, at least 25% of the European *CFTR*-testing laboratories do not participate in the CF EQA scheme at all. Some of these are likely to participate in national CF EQA schemes organised by UKNEQAS, IEQA-ISS, Afssaps or INSTAND (only for c.1521_1523delCTT).^{20–23} We expect similar percentages in schemes for other disorders or methods. It seems likely that laboratories that do not participate in EQA and whose reports have never been objectively assessed by peers and against expert opinions would have relatively low interpretation scores, comparable to those that have participated in only 1 of the 3 investigated years (between 1.2/2.0 and 1.7/2.0). Therefore, participation of laboratories in EQA on a regular basis, for all disorders they test for, needs to be continuously encouraged. The adoption of ISO 15189, for implementing a quality management system and assessing technical competence, in national legislations could accelerate the implementation of accreditation, and thus assure EQA participation in genetic testing laboratories. Today, this is the case in only a few European countries,

for example in France where accreditation will be mandatory by 2016 for medical biological laboratories, and in Switzerland where EQA participation for all genetic tests is becoming mandatory.^{24,25} In the United States, the situation is different and all laboratories conducting moderate and/or high complexity testing are required by law (CLIA) to participate in PT for certain tests they perform. However, the CLIA regulations do not have PT requirements specific for molecular genetic tests. Therefore, laboratories that perform genetic tests must comply with the general requirements for alternative performance assessment for any test or analyte not specified as a regulated analyte to verify the accuracy of any genetic test or procedure they perform.²⁶ Unfortunately in the United States, focus is mainly on the assessment of the accuracy of the test result (genotyping) and not on interpretation assessment, which is sufficient to comply with CLIA.⁵ We conclude that EQA contributes to improved interpretation of sensitive genetic test results and improved interpretation will lead to better quality of care for patients and their families by all healthcare professionals involved in and collaborating with genetic testing services.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank Dr Annoushka Laenen, from the Leuven Biostatistics and Statistical Bioinformatics Centre, for carrying out the statistical analysis. We would like to acknowledge Nick Nagels and Romy Gentens for providing data management and administrative assistance. We are grateful to all the laboratories that participated in the CF EQA schemes over the years and their commitment to provide high quality in genetic testing. This study was funded by the EU project EuroGentest (FP6-512148).

- 1 Stenhouse SA, Middleton-Price H: Quality assurance in molecular diagnosis: the UK experience. *Methods Mol Med* 1996; **5**: 341–352.
- 2 Organisation for Economic Co-operation and Development: OECD guidelines for quality assurance in molecular genetic testing 2007.
- 3 International Organization for Standardization: ISO 15189:2007 Medical laboratories - Particular requirements for quality and competence 2007.
- 4 International Organization for Standardization: ISO/IEC 17025:2005 General requirements for the competence of testing and calibration laboratories 2005.
- 5 Clinical Laboratory Improvement Amendments (CLIA): <http://www.cdc.gov/clia/regs/toc.aspx>, 2011.
- 6 Cystic Fibrosis European Network. <http://cf.eqascheme.org/>, 2011.
- 7 Toutou I, Rittore C, Philibert L *et al*: An international external quality assessment for molecular diagnosis of hereditary recurrent fevers: a 3-year scheme demonstrates the need for improvement. *Eur J Hum Genet* 2009; **17**: 890–896.
- 8 Seneca S, Morris MA, Patton S *et al*: Experience and outcome of 3 years of a European EQA scheme for genetic testing of the spinocerebellar ataxias. *Eur J Hum Genet* 2008; **16**: 913–920.
- 9 Mueller CR, Kristofferson U, Stoppa-Lyonnet D: External quality assessment for mutation detection in the BRCA1 and BRCA2 genes: EMQN's experience of 3 years. *Ann Oncol* 2004; **15**(Suppl 1): I14–I17.
- 10 McGovern MM, Elles R, Beretta I *et al*: Report of an international survey of molecular genetic testing laboratories. *Community Genet* 2007; **10**: 123–131.
- 11 Ramsden SC, Deans Z, Robinson DO *et al*: Monitoring standards for molecular genetic testing in the United Kingdom, the Netherlands, and Ireland. *Genet Test* 2006; **10**: 147–156.
- 12 Qiu J, Hutter P, Rahner N *et al*: The educational role of external quality assessment in genetic testing: a 7-year experience of the European Molecular Genetics Quality Network (EMQN) in Lynch syndrome. *Hum Mutat* 2011; **32**: 696–697.
- 13 Technical Standards and Guidelines for CFTR. Mutation Testing: http://www.acmg.net/Pages/ACMG_Activities/stds-2002/cf.htm, 2011.
- 14 Castellani C, Cuppens H, Macek Jr. M *et al*: Consensus on the use and interpretation of cystic fibrosis mutation analysis in clinical practice. *J Cyst Fibros* 2008; **7**: 179–196.
- 15 Dequeker E, Stuhmann M, Morris MA *et al*: Best practice guidelines for molecular genetic diagnosis of cystic fibrosis and CFTR-related disorders - updated European recommendations. *Eur J Hum Genet* 2009; **17**: 51–65.
- 16 Schwarz M, Gardner A, Jenkins L, Norbury G, Renwick P, Robinson D: *Testing Guidelines for Molecular Diagnosis of Cystic Fibrosis*. United Kingdom: Clinical Molecular Genetics Society, 2009.
- 17 Orphanet. <http://www.orpha.net/>, 2011.

- 18 Cassiman JJ: Research network: EuroGentest - a European Network of Excellence aimed at harmonizing genetic testing services. *Eur J Hum Genet* 2005; **13**: 1103–1105.
- 19 Berwouts S, Fanning K, Morris MA, Barton DE, Dequeker E: Quality assurance practices in Europe: a survey of molecular genetic testing laboratories. *Eur J Hum Genet* 2012; e-pub ahead of print 20 June 2012; doi:10.1038/ejhg.2012.125.
- 20 UKNEQAS. <http://www.ukneqas-molgen.org.uk/ukneqas/index/news.html>, 2011.
- 21 Salvatore M, Falbo V, Florida G *et al*: The Italian External Quality Control Programme for cystic fibrosis molecular diagnosis: 4 years of activity. *Clin Chem Lab Med* 2007; **45**: 254–260.
- 22 INSTAND e.V. <http://www.instandev.de/en/eqas/programm/>, 2011.
- 23 Agence Francaise de Securite Sanitaire des Produits de Sante (afssaps). <http://www.afssaps.fr/Activites/Controle-national-de-qualite-des-analyses-de-biologie-medica-CNQ>, 2011.
- 24 La loi Hôpital, Patients, Santé et Territoires (HPST). <http://www.sante.gouv.fr/la-loi-hopital-patients-sante-et-territoires.html>, 2011.
- 25 Office fédéral de la santé publique (OFSP). <http://www.bag.admin.ch/themen/medizin/00683/02724/03677/11167/index.html?lang=fr>, 2011.
- 26 Chen B, Gagnon M, Shahangian S *et al*: Good laboratory practices for molecular genetic testing for heritable diseases and conditions. *MMWR Recomm Rep* 2009; **58**: 1–37.