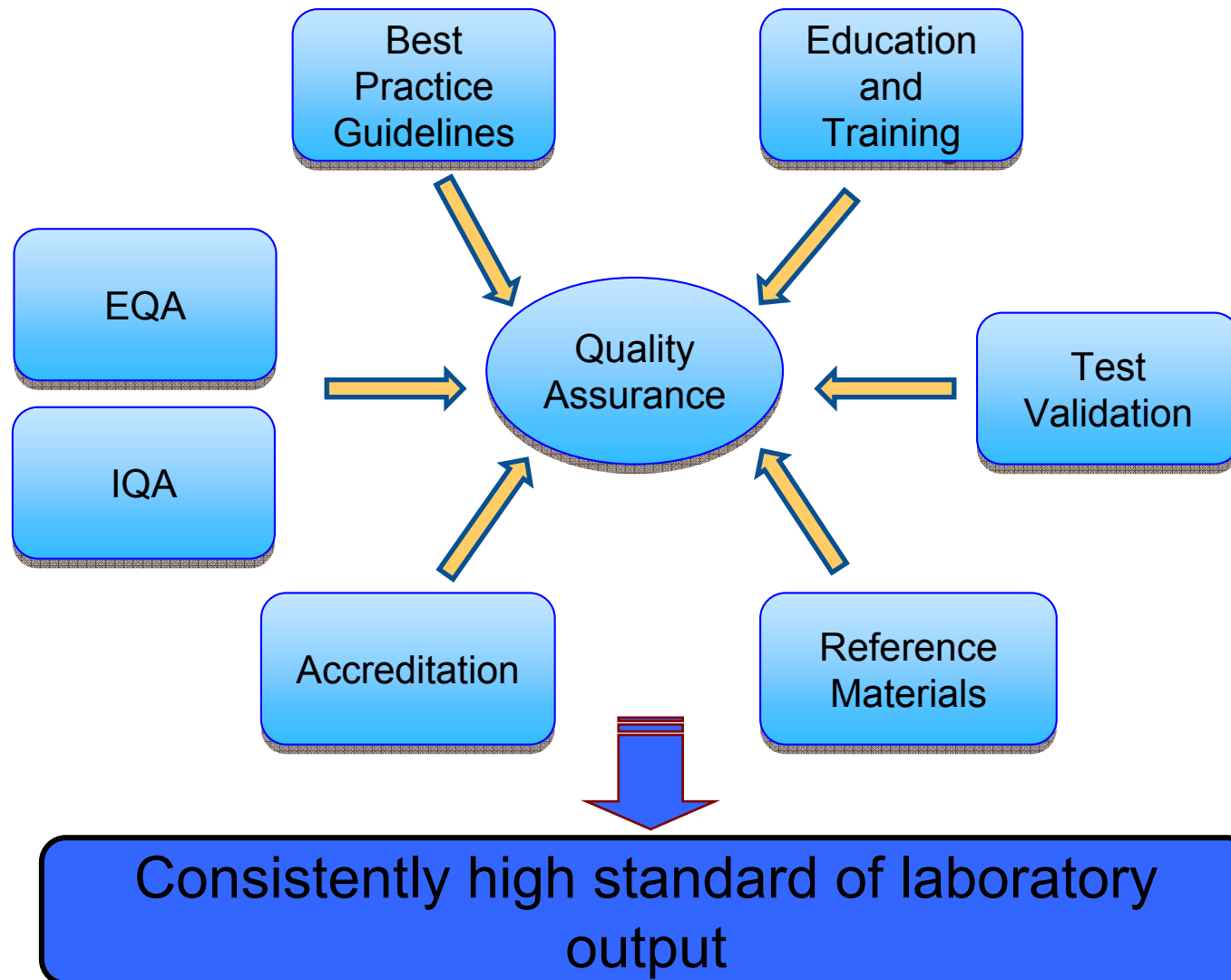


Quality in the World of Molecular Genetics

David Barton, EMQN
National Centre for Medical Genetics
Dublin, Ireland

Quality Assured Genetic Testing



Molecular Genetics Testing



- Mostly yes / no answers
- TAT 3 to 40 days
- Scientists give the clinical interpretation of results
- Permanent medical record

Current Molecular Genetics Testing

Mendelian disorders

- Cancer
- Neurological disorders
- Developmental delay
- Dysmorphology
- Etc.

Chromosomal
imbalance

Molecular pathology
KRAS, GIST

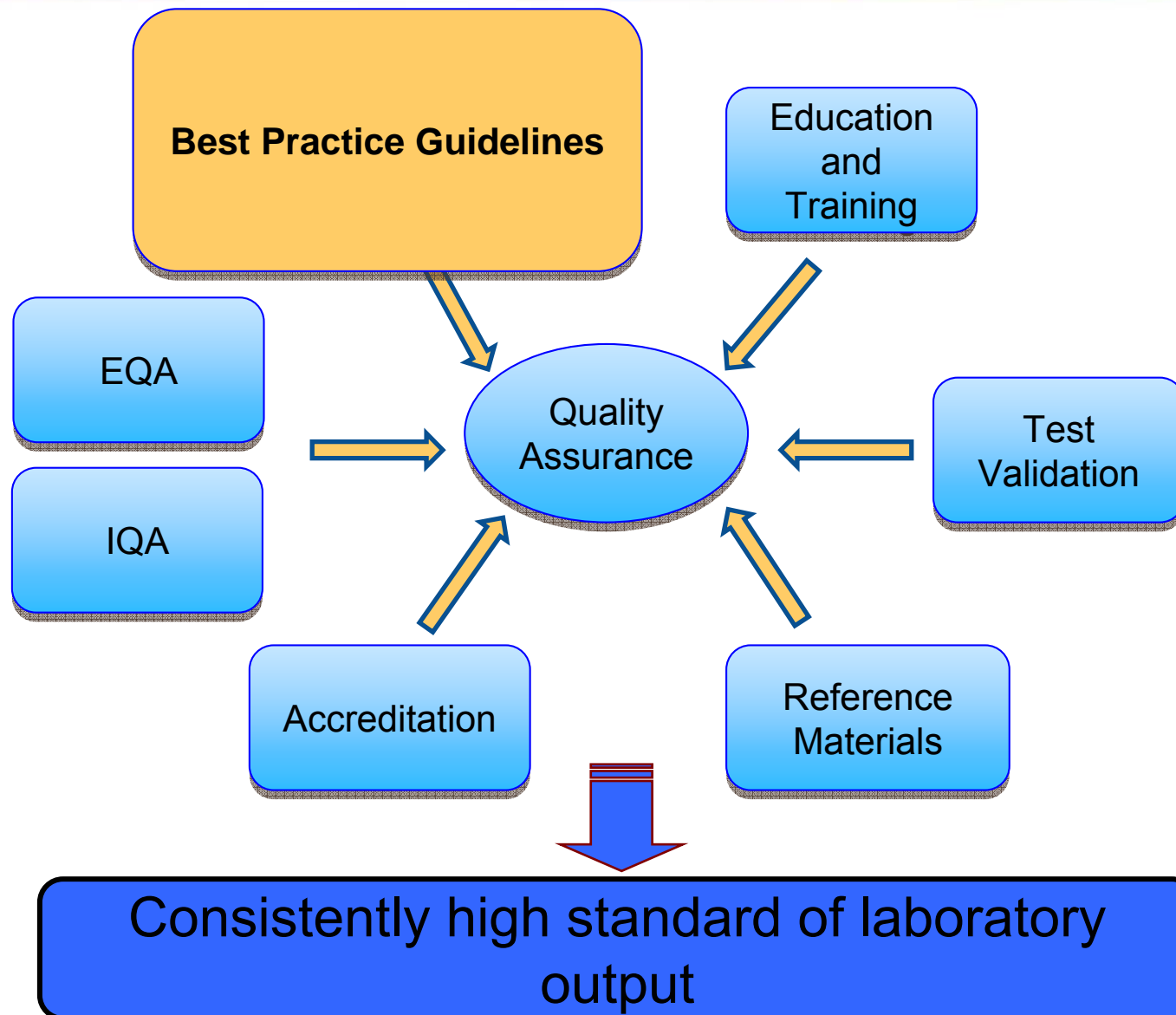
Non-invasive
prenatal diagnosis

Future directions

Free fetal DNA analysis

Next generation sequencing
multi-gene disorders

Pharmacogenetics



Best practice

- Sequencing
- Single nucleotide polymorphisms
- Unclassified variants – pathogenic or not?

Best practice

- Sequencing
 - Uni – or bi-directional?
 - Genotyping
 - Confirmation or exclusion of known sequence variant
 - Mutation scanning
 - Full characterisation of a region of DNA

TACTTAG **CMGS** CGTGTTCAGTACCGTACTTAGCGT
CGTACT CGTGTTCAGTACCGTACTCGTGT
TCAGTACCGTACTTAG **CLINICAL MOLECULAR** ACGTGTC
GTGTCAGTACCGTACT **GENETICS SOCIETY** ACGTGTACG

Practice guidelines for Sanger Sequencing Analysis and Interpretation.

Prepared and edited by Sian Ellard¹, Ruth Charlton², Michael Yau³, David Gokhale⁴, Graham R Taylor⁵, Andrew Wallace⁶ and Simon C Ramden⁶. Ratified by the CMGS Executive Committee on 7th August, 2009.

1. Department of Molecular Genetics, Royal Devon & Exeter NHS Foundation Trust, Barrack Road, Exeter, EX2 5AD, UK.
2. Yorkshire Regional DNA Laboratory, St James's University Hospital, Leeds LS9 7TF, UK.
3. Guy's & St Thomas's Hospital NHS Trust, Genetics Centre, 8th Floor, Guy's Tower, Guy's Hospital, London SE1 9RT, UK.
4. Liverpool Women's NHS Foundation Trust, Crown Street, Liverpool, L8 7SS, UK.
5. Genomics Facility, Leeds Institute of Molecular Medicine & Leeds Teaching Hospitals, 6.2 Clinical Sciences Building, St James's University Hospital, Leeds, LS9 7TF, UK.
6. Regional Genetic Service, St Mary's Hospital, Hathersage Road, Manchester, M13 0JH, UK.

1. INTRODUCTION

1.1 General Introduction

DNA sequencing is the most commonly used approach for both mutation scanning and mutation testing; it is widely regarded as the gold standard. Agreed practice guidelines for both the sequencing process and the interpretation of results are important to achieve a high quality approach with common quality standards across different laboratories. These guidelines do not constitute an experimental protocol or troubleshooting guide, rather they aim to establish consensus standards for identifying and reporting mutations.

Different standards will be required for clinical diagnostics than would be acceptable for a sequence-based research project. Since germline changes are most frequently being analysed, results will stand for the lifetime of the individual and may have implications for relatives of the proband.

This document considers quality aspects of the whole process of sequencing and makes the assumption that the analytical process takes place in an appropriate, accredited laboratory setting where routine aspects of good laboratory practice such as sample tracking and record keeping are in place.

Local sequencing practices may vary both in terms of reasons for investigation, chemistry, hardware, software and reporting of results. These guidelines have been updated from an earlier version by Ravine *et al* (no longer available) that followed a CMGS Sequencing Best Practice Meeting held in 2001. These guidelines identify common elements for each part of the process and specify quality criteria that should be met or exceeded.

Guidelines are described as either:

- **Essential** practice which must be implemented to ensure quality of service
- **Recommended** practice where more than one approach is satisfactory, however there is a clear advantage in following the advice given.

- *Not acceptable*, which highlights areas where the quality of service may be compromised.

1.2 Reasons for diagnostic sequencing

The reason for the sequencing investigation may influence the quality standards required. Two types of investigation can be considered:

- 1) Confirmation or exclusion of a known sequence variant (genotyping).

- 2) Full characterisation of a defined region of DNA (mutation scanning or re-sequencing).

From the perspective of the required quality standards these two processes are not equivalent; whereas genotyping is concerned only with base changes previously identified, fully characterising a region of DNA implies that each base has been determined with high confidence. So if the average probability of error for each base call of a 1 kilobase region was 1% (equivalent to a Phred score of 20 – see later), then one would expect 10 erroneous base calls which is clearly unacceptable for diagnostic purposes.

2. QUALITY ASPECTS OF THE LABORATORY PROCESS

2.1 Patient material

Suitable material for sequence analysis is essential. This requires correct identification of the proband, appropriate clinical diagnosis and the sample must be collected, identified, recorded and stored under quality controlled conditions appropriate for diagnostic testing. For example if a case has been identified as part of a research project it may be necessary to collect an additional sample. Genomic DNA from peripheral white blood cells is the typical starting material. Alternative sources such as fixed tissue or cDNA may raise quality control issues that are beyond the scope of these guidelines.

Best practice

- Sequencing
 - Quality parameters not standardized
 - Negative results

TACTTAG **CMGS** CGTGTTCAGTACCGTACTTAGCGT
:CGTACT :ACGTGTTCAGTACCGTACTCGTGT
TCAGTACCGTACTTAG **CLINICAL MOLECULAR** ACGTGTCTC
GTGTTCAGTACCGTACT **GENETICS SOCIETY** ACGTGTTCAGT

Practice guidelines for Sanger Sequencing Analysis and Interpretation.

Prepared and edited by Sian Ellard¹, Ruth Charlton², Michael Yau³, David Gokhale⁴, Graham R Taylor⁵, Andrew Wallace⁶ and Simon C Ramden⁶. Ratified by the CMGS Executive Committee on 7th August, 2009.

1. Department of Molecular Genetics, Royal Devon & Exeter NHS Foundation Trust, Barrack Road, Exeter, EX2 5AD, UK.
2. Yorkshire Regional DNA Laboratory, St James's University Hospital, Leeds LS9 7TF, UK.
3. Guy's & St Thomas's Hospital NHS Trust, Genetics Centre, 8th Floor, Guy's Tower, Guy's Hospital, London SE1 9RT, UK.
4. Liverpool Women's NHS Foundation Trust, Crown Street, Liverpool, L8 7SS, UK.
5. Genomics Facility, Leeds Institute of Molecular Medicine & Leeds Teaching Hospitals, 6.2 Clinical Sciences Building, St James's University Hospital, Leeds, LS9 7TF, UK.
6. Regional Genetic Service, St Mary's Hospital, Hathersage Road, Manchester, M13 0JH, UK.

1. INTRODUCTION

1.1 General Introduction
DNA sequencing is the most commonly used approach for both mutation scanning and mutation testing; it is widely regarded as the gold standard. Agreed practice guidelines for both the sequencing process and the interpretation of results are important to achieve a high quality approach with common quality standards across different laboratories. These guidelines do not constitute an experimental protocol or troubleshooting guide, rather they aim to establish consensus standards for identifying and reporting mutations.

Different standards will be required for clinical diagnostics than would be acceptable for a sequence-based research project. Since germline changes are most frequently being analysed, results will stand for the lifetime of the individual and may have implications for relatives of the proband. This document considers quality aspects of the whole process of sequencing and makes the assumption that the analytical process takes place in an appropriate, accredited laboratory setting where routine aspects of good laboratory practice such as sample tracking and record keeping are in place.

Local sequencing practices may vary both in terms of reasons for investigation, chemistry, hardware, software and reporting of results. These guidelines have been updated from an earlier version by Ravine *et al* (no longer available) that followed a CMGS Sequencing Best Practice Meeting held in 2001. These guidelines identify common elements for each part of the process and specify quality criteria that should be met or exceeded.

Guidelines are described as either:

- **Essential** practice which must be implemented to ensure quality of service
- **Recommended** practice where more than one approach is satisfactory, however there is a clear advantage in following the advice given.

- *Not acceptable*, which highlights areas where the quality of service may be compromised.

1.2 Reasons for diagnostic sequencing

The reason for the sequencing investigation may influence the quality standards required. Two types of investigation can be considered:

- 1) Confirmation or exclusion of a known sequence variant (genotyping).
- 2) Full characterisation of a defined region of DNA (mutation scanning or re-sequencing).

From the perspective of the required quality standards these two processes are not equivalent; whereas genotyping is concerned only with base changes previously identified, fully characterising a region of DNA implies that each base has been determined with high confidence. So if the average probability of error for each base call of a 1 kilobase region was 1% (equivalent to a Phred score of 20 – see later), then one would expect 10 erroneous base calls which is clearly unacceptable for diagnostic purposes.

2. QUALITY ASPECTS OF THE LABORATORY PROCESS

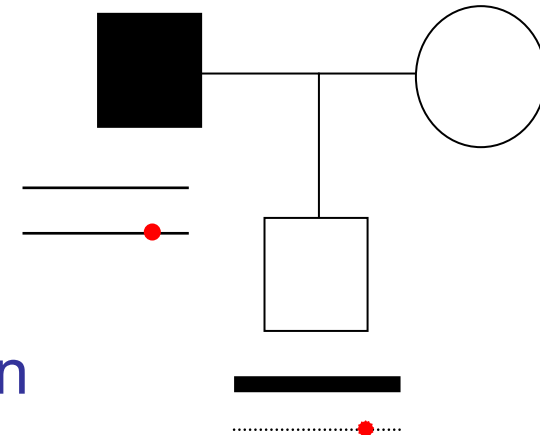
2.1 Patient material

Suitable material for sequence analysis is essential. This requires correct identification of the proband, appropriate clinical diagnosis and the sample must be collected, identified, recorded and stored under quality controlled conditions appropriate for diagnostic testing. For example if a case has been identified as part of a research project it may be necessary to collect an additional sample. Genomic DNA from peripheral white blood cells is the typical starting material. Alternative sources such as fixed tissue or cDNA may raise quality control issues that are beyond the scope of these guidelines.

SNP check

Case:

- HNPCC family
 - Father affected – known mutation
 - Test result was negative for the son
-
- Need to check if there are any known SNPs under primer binding sites - SNPcheck



EQA genotyping results

	Genotyping errors per case			Genotyping errors per allele		
	No. of cases completed	No. of errors	Error rate (%)	No. of alleles analysed	No. of errors	Error Rate (%)
Case 1	45	0	0.0	90	0	0.0
Case 2	47	10	21.2	94	10	10.6

Re-design primers

- Move primer binding site
- Use wobble primers

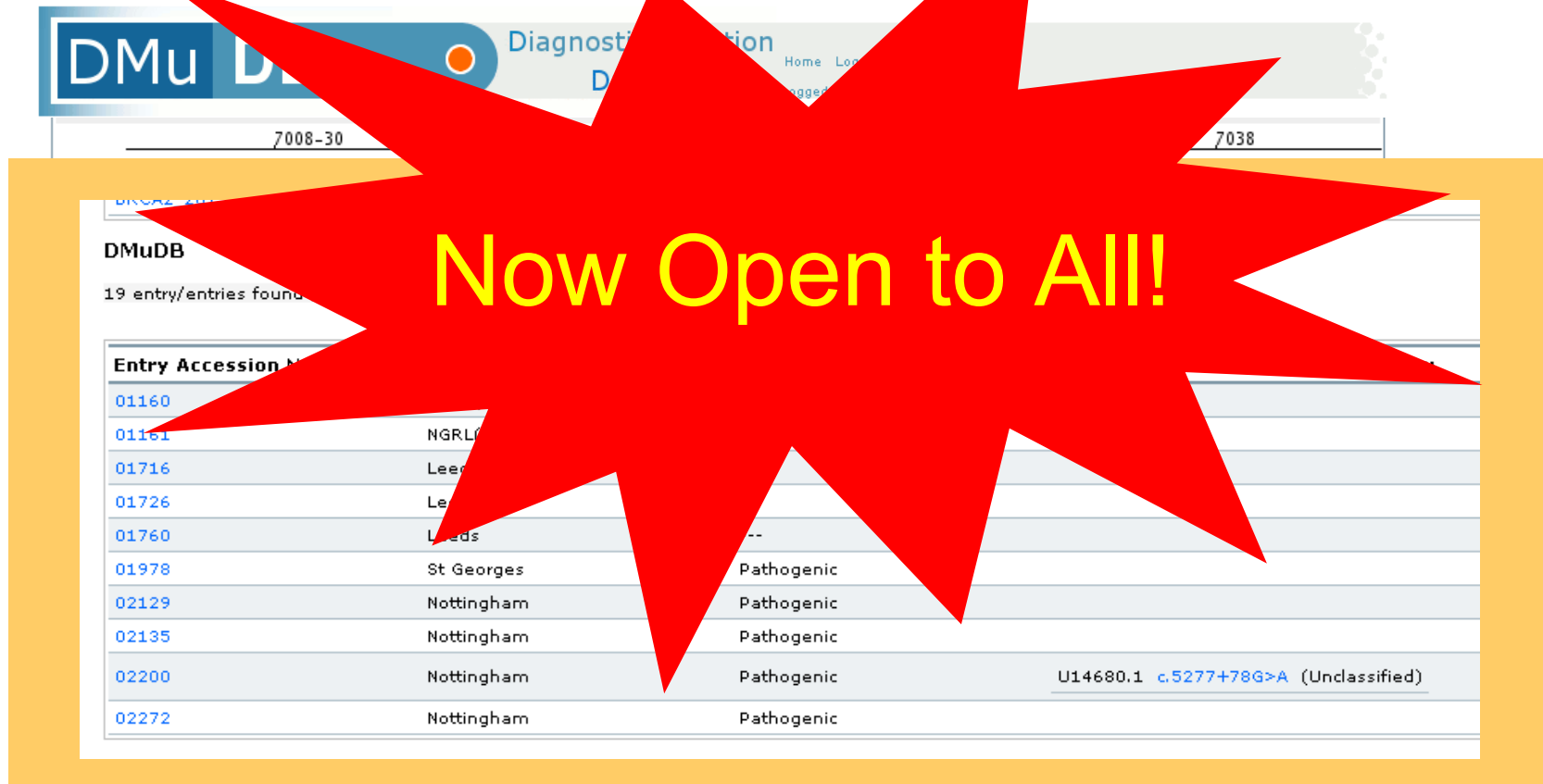
Diagnostic error rate

Best practice

- Unclassified variants
 - Mutation or normal variant?
- Guidelines published
- Bioinformatics tools
 - Training/experience/caution required
 - Very time-consuming
- Novel approaches
 - Alamut
 - Bayesian Classification (Tavtigian, BRCA)

Best practice

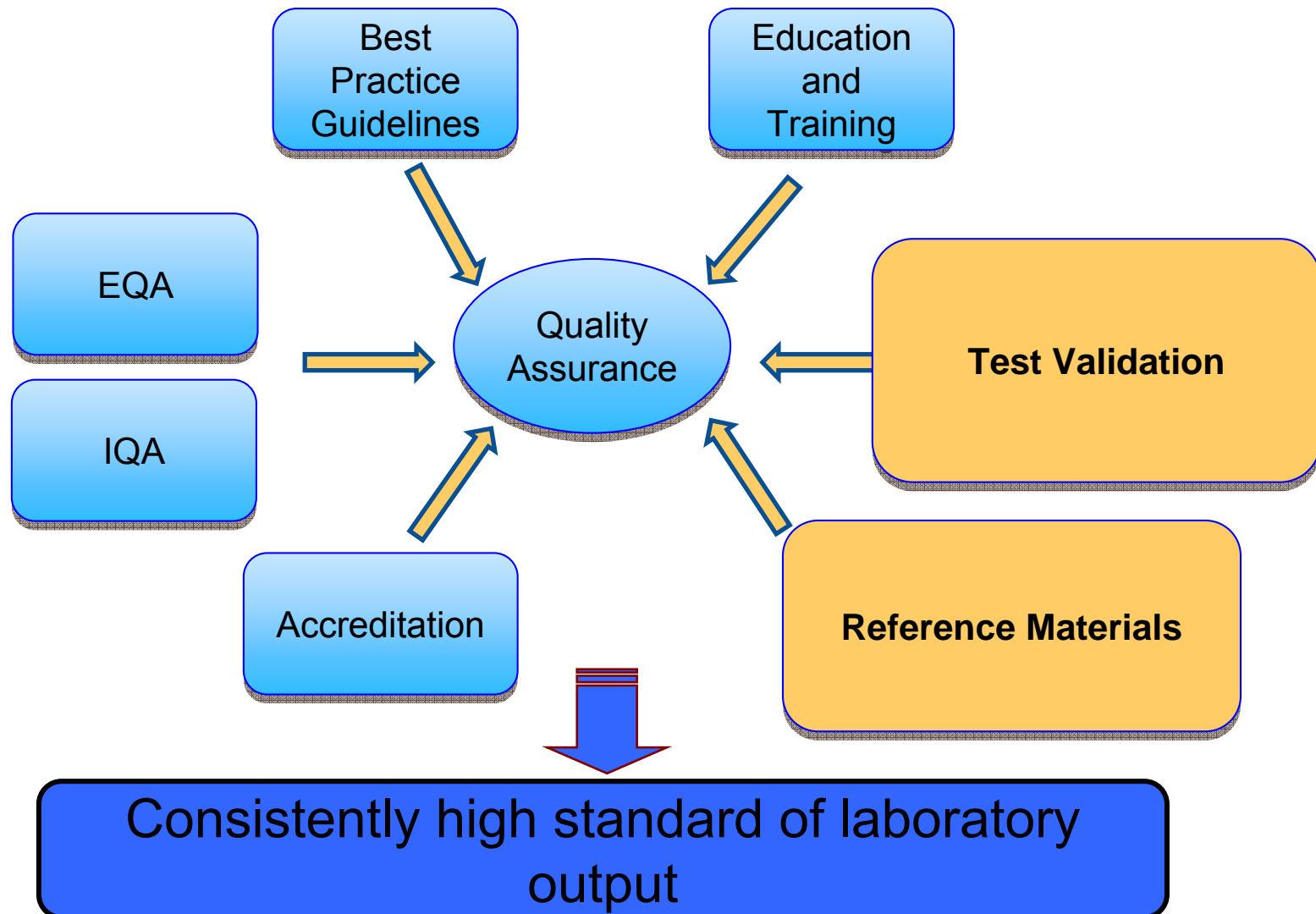
- Unclassified variants
 - DMuDB – Diagnostic Mutation Database



The screenshot shows the DMuDB (Diagnostic Mutation Database) website. A large red starburst with the text "Now Open to All!" is overlaid on the page. The website header includes the DMuDB logo and navigation links. Below the header, there is a search bar and a table of entries. The table has columns for Entry Accession Number, Institution, and Pathogenicity. The table lists 19 entries, with the last entry being U14680.1 c.5277+78G>A (Unclassified).

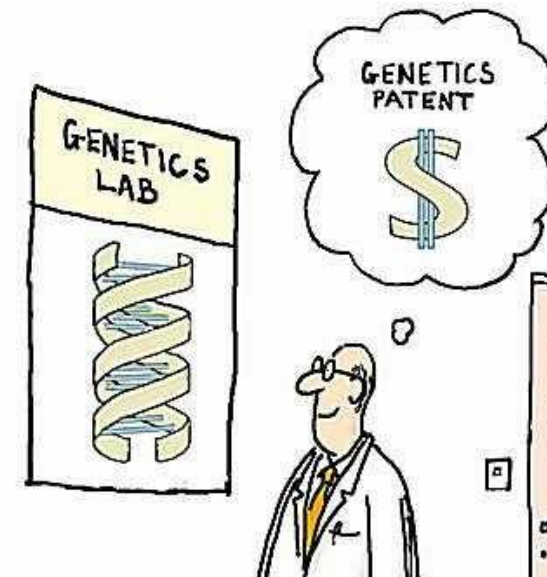
Entry Accession Number	Institution	Pathogenicity
01160		
01161	NGRLC	
01716	Leeds	
01726	Leeds	
01760	Leeds	
01978	St Georges	Pathogenic
02129	Nottingham	Pathogenic
02135	Nottingham	Pathogenic
02200	Nottingham	Pathogenic
02272	Nottingham	Pathogenic
U14680.1	c.5277+78G>A	(Unclassified)

Quality Assured Genetic Testing

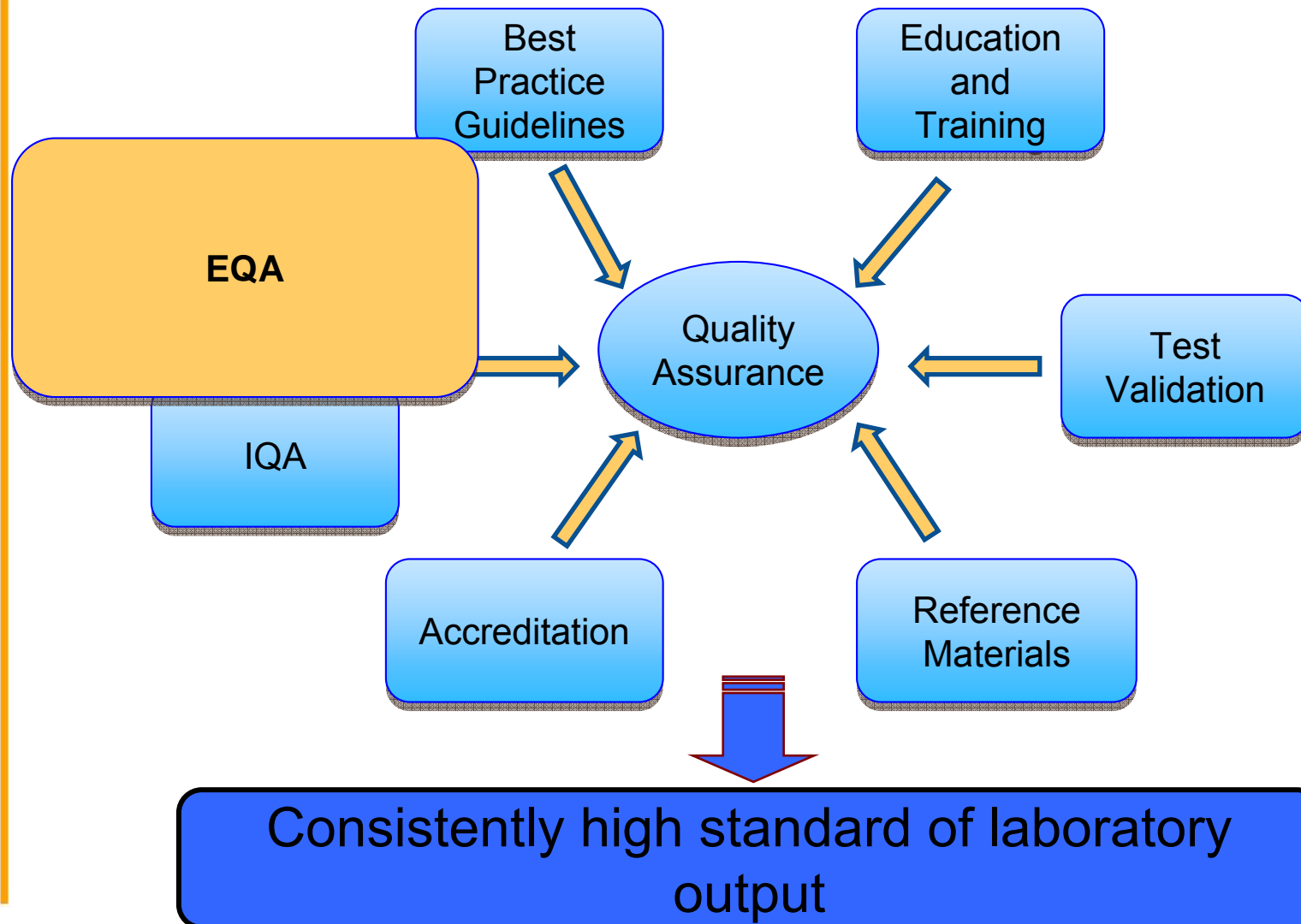


Test validation

- What is a new test?
- Most tests used are 'homebrews'
- Lack of reference material
 - Panel for FRAX 2008
 - **Panel for PWS/AS**
 - **Panel for BCR-ABL Q-PCR**

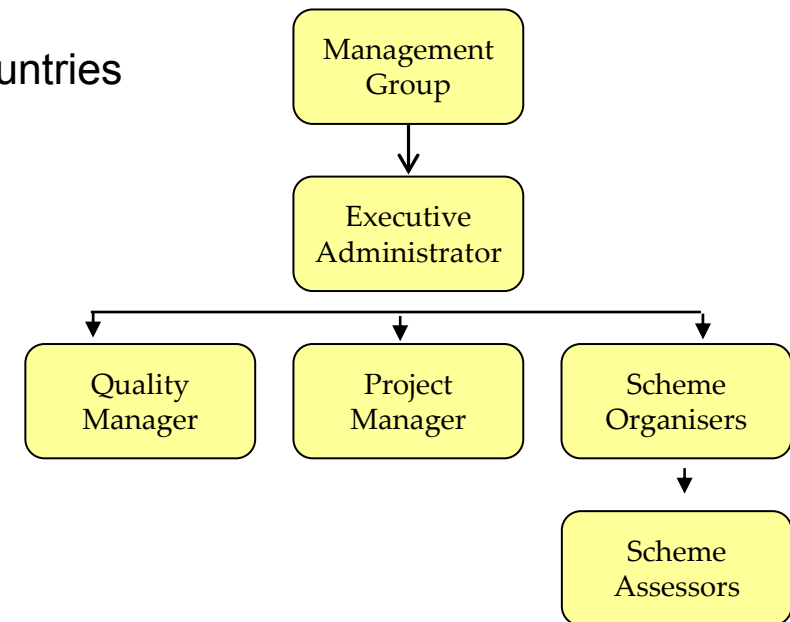
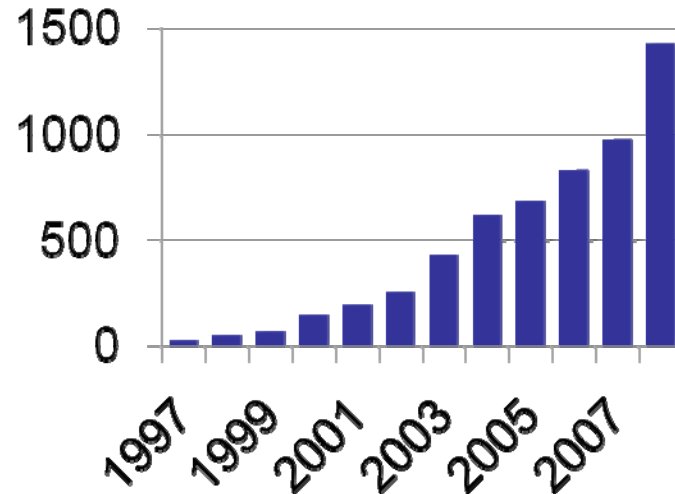


Quality Assured Genetic Testing



EMQN – network at a glance

- Manchester based international EQA organiser
- Network started in 1997 (EU funding)
- A management group of **9** scientists from **7** countries
- **52** scheme organisers and assessors from 11 countries
- **586** registered members
- **1400** participations



Scheme participation

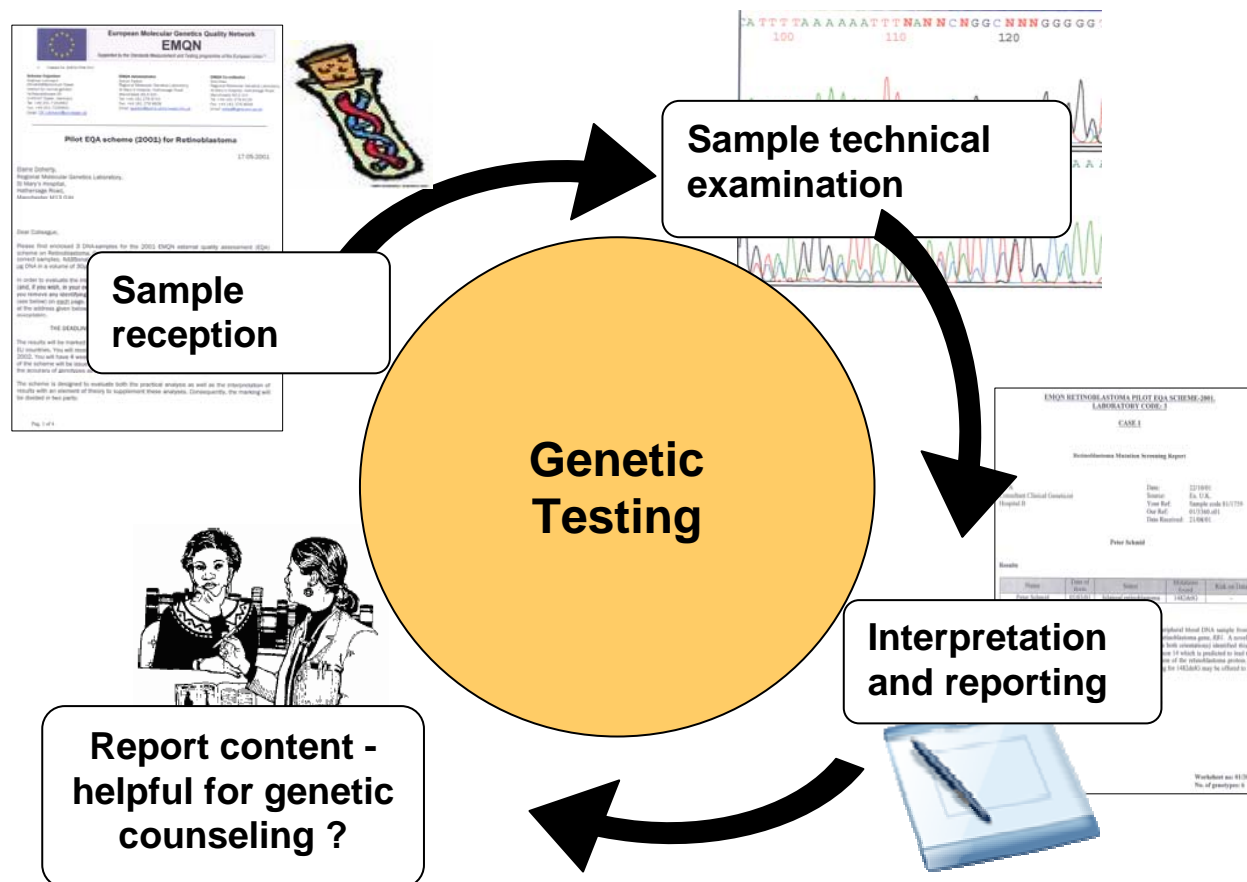
Scheme	2004	2005	2006	2007	2008	Difference
AZF	91	88	97	96	125	+29
BRCA	51	58	70	77	101	+34
CMT	37	37	42	50	73	+12
CAH	0	0	0	0	38	+38
CNX26	0	0	0	0	30	+30
DM	0	0	0	31	73	+42
DMD	33	42	41	46	58	+12
FRAX	54	55	79	75	102	+27
FRDA	26	33	33	34	52	+18
HD	39	45	59	57	87	+30
HFE	58	55	60	61	67	+7
HNPCC	42	49	51	63	81	+18
Men2	0	0	0	0	44	+44
MonoDiab	0	0	12	13	20	+7
PKU	11	9	14	17	15	-2
POR	0	9	11	16	15	-1
PWAS	44	47	47	49	82	+33
RB	10	10	11	14	14	0
SCA	32	37	28	34	62	28
SMA	0	0	32	57	73	+16
VHL	0	0	0	0	20	+20
WIL	0	0	22	23	23	0
SEQ	95	114	122	127	136	9
MSCAN	0	0	0	41	36	-5
Total:	623	688	831	981	1427	+446 (+45%)

International schemes



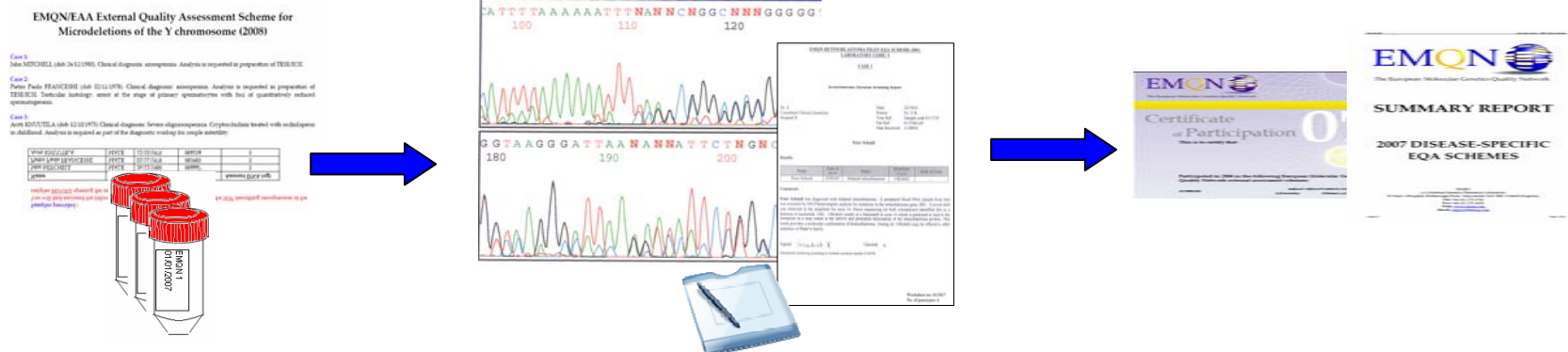
Types of schemes

- Single gene disorder schemes
- Technique specific schemes
 - Qualitative
 - Quantitative



Disease specific EQA schemes

- Participants receive 3 samples once a year with mock clinical referral
- Laboratories carry out analysis and report findings
- Assessment for genotyping, interpretation and reporting
- Participants receive their marks and a final report



Scheme assessment

- Genotyping – correct?
- Reporting
 - Accuracy and consistency
 - Clear ‘take home’ message
 - Methods referenced
 - Authorised / audit trail
- Interpretation
 - Can be scored
 - Language can be accommodated
 - Reporting policies differ

From lab code: _____ To: Dr. PAEDIATRICIAN
December 10th 2007

EMQN
Report Myotonic Dystrophy European EQA, Scheme 2007
Diagnostic testing for Myotonic Dystrophy, type 1

QA: ---
Consent: ---
URRef: case 1, batch 075302, diagnostic testing

Patient identification: Camilla LUNDBY, B12.03.2007
Requested by: Dr. Paediatrician (EMQ EQA, scheme 2007)
Date of request: 24.10.2007 Date of result: 10.12.2007
Lab identification number: 07.3738 Sex: female
Material: DNA sample extracted from lymphoblastoid cell lines

Clinical information: Camilla LUNDBY died few months of age. Polyhydramnios, decreased fetal movements were observed in the pregnancy. She had respiratory failure and clubfoot, was hypotonic. DNA analysis for SMA was normal. Also normal results from a metabolic screening. Clinical suspicion of congenital DM1. No family history.

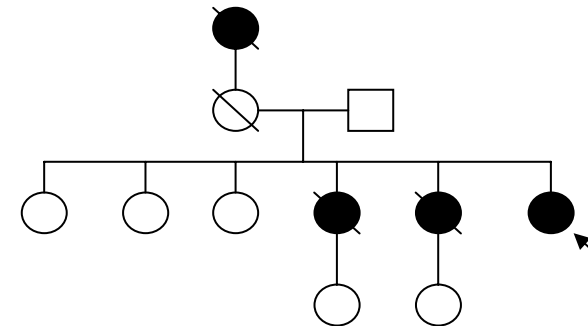
Result:
allele 1: 12 repeats
allele 2: around 650-700 repeats

Conclusion:
Diagnostic testing for myotonic dystrophy type 1 mutations indicates that the patient has an abnormal expanded allele in the *DMPK* gene. The clinical suspicion is confirmed. Camilla Lundby was affected with congenital myotonic dystrophy, type 1. This diagnosis has also implications for other family members at risk. These results implicate that the mother of Camilla Lundby is most probably affected with DM, while siblings of the patient have also a 50% risk of being affected. Molecular testing of the parents is strongly recommended. A counseling session at the Center for Medical Genetics is necessary. If appropriate, prenatal diagnosis is possible.

Methodology:
Determination of the CTG repeat number in the 3' untranslated region of the *DMPK* gene on chromosome 19q13 by PCR and TP-PCR strategy as described by Bosio et al. (1992) Cell 68: 799-803, and Warner et al. (1996) J Med Genet 33:1022-1026, and by Southern blot analysis. The procedure was adapted for the ABI3130 detection system. The accuracy of sizing is ± 1 repeat for the normal allele size range.
Number of repeats in normal person: 5-13
Number of repeats in affected patients: 50 and more

Molecular Geneticist, Ph.D. Clinical Geneticist, MD.

EMQN BRCA scheme case 2007



Claudia CAPABLANCA (dob 21/08/1976) is the youngest of 6 sisters. Two of her older sisters and her grand-mother all died before menopause from breast or ovarian cancers. Her mother had died in an accident at age 49. Mrs. Capablanca's gynaecologist recently suspected an ovarian cancer and referred her to an oncology centre. There, the diagnosis was confirmed and treatment initiated. Because of the high genetic risk for her nieces, Mrs. Capablanca decided to have a genetic screening done.

Please analyse exon 18 of the BRCA1 gene. Report back to the oncology centre in your standard reporting format.

EQA evaluation of results

Name	Sex	Date of Birth	Results
Claudia CAPABLANCA	FEMALE	21/08/1976	U14680.1: c.5075-2A>C

Name	Criteria	Marks
Genotyping	Genotype correct , deduction of marks only if genotype description is misleading.	2.0
Interpretation and reporting	<p>Mutation affects 5'-splice site of exon 18. Listed once in BIC as 'clinically relevant'. Most likely compatible with a hereditary predisposition to breast/ovarian cancer.</p> <p>Patient remains at high risk for secondary tumours and should join a monitoring programme. Predictive testing of relatives is possible after genetic counselling.</p> <p><u>Deductions:</u> missing/incomplete biological interpretation; lack of suggestion of monitoring programme (0.4 marks); lack of suggestion of predictive testing and/or counselling (0.2 marks).</p> <p>Personal data of patient (name & DOB or name & lab no) Brief recapitulation of the patient's personal & family history Clear presentation of the results <i>Minor points (not leading to deduction of marks):</i> Arrival and reporting dates, Signature of the report by two authorised persons, Indication of the reference sequence used. <u>Deductions:</u> (0.2 marks each)</p>	2.0
Comments	BIC-nomenclature: c.5194-2A>C; IVS17-2A>C	

Technique specific EQA

- EQUAL – qual
 - EQA for DNA extraction and PCR
- EQUAL – quant
 - EQA for Real Time PCR
- EMQN schemes
 - Sequencing
 - Mutation scanning
 - CNVs/array CGH

- EQA for DNA extraction and PCR
 - Participants received blood, primers and DNA
 - Labs asked to extract DNA and set up PCR reactions
 - Submit DNA concentration, quality, quantity data as well as aliquots of DNA and PCR products

EQUAL – qual results

- Assessment:
 - Look at data returned by participants
 - Reevaluation of returned material by reference laboratory
 - Assessment of photometric measurements of DNA quantity / quality
 - Performance of blood extraction procedure

EQUAL - qual

Table 1. Simple data description of DNA quantity and quality in blood and pre-extracted samples (preamplification phase).^a

Variable	n	Minimum	25th centile	Median	75th centile	Maximum	IQR
R1	166	0.36	1.65	1.78	1.89	5.14	0.23
R2	166	0.50	1.65	1.78	1.90	8.04	0.25
R3	164	0.29	1.66	1.81	1.95	18.00	0.29
R4	165	0.91	1.33	1.38	1.46	5.00	0.13
Q1	150	1.67	11.50	17.00	22.33	187.50	10.83
Q2	152	1.67	13.32	19.69	25.29	200.00	11.97
Q3	148	5.50	17.86	20.13	30.00	200.00	12.14
Q4	148	20.00	49.98	54.25	65.65	190.00	15.68
Re1	166	0.90	1.69	1.84	1.90	13.71	0.21
Re2	166	0.52	1.64	1.83	1.90	2.94	0.25
Qe1	149	1.42	9.13	14.41	19.95	127.63	10.82
Qe2	151	0.48	11.52	17.85	25.25	361.10	13.73
$\Delta R1$	166	-3.29	-0.15	0.03	0.19	13.21	0.34
$\Delta R2$	166	-7.04	-0.21	0.03	0.19	1.43	0.39
$\Delta Q1$	152	-86.00	-6.38	-1.22	0.94	51.35	7.32
$\Delta Q2$	148	-105.10	-5.25	-0.70	1.87	94.10	7.12

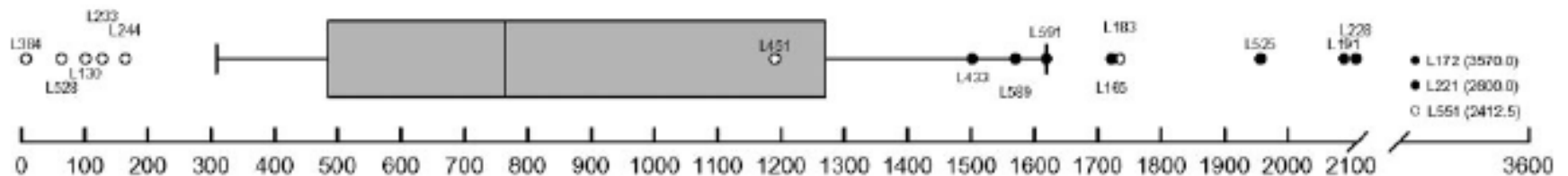
^a R1, R2, R3, and R4 are the (260 nm – 320 nm)/(280 nm – 320 nm) or 260 nm/280 nm ratio for samples 1, 2, 3, and 4, respectively. Q1, Q2, Q3, and Q4 are the quantity of DNA extracted from blood samples 1 and 2 (mg/L) obtained from [(260 nm – 320 nm) × 50 × dilution factor × DNA reconstitution volume]/extracted blood volume] or directly measured in pre-extracted samples 3 and 4, respectively. Re1 and Re2 are the (260 nm – 320 nm)/(280 nm – 320 nm) ratios measured in the EQUAL-Laboratory by Nanodrop in blood samples 1 and 2, respectively, as provided by participants. Qe1 and Qe2 are the DNA quantities measured in the EQUAL-Laboratory with Nanodrop in blood samples 1 and 2, respectively, as provided by participants. $\Delta R1 = Re1 - R1$. $\Delta R2 = Re2 - R2$. $\Delta Q1 = Qe1 - Q1$. $\Delta Q2 = Qe2 - Q2$.

- 25 % of laboratories (42/165) gave out of limits readings for pre-extracted DNA

EQUAL - qual

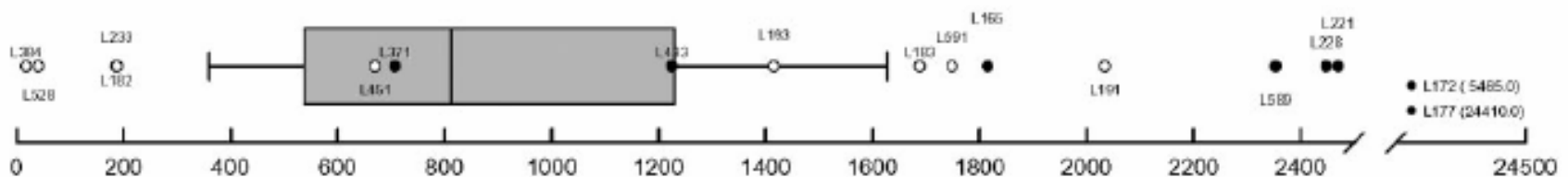
PCR from DNA extracted by the participants

Panel A: E1



PCR from DNA extracted by the organiser

Panel A: E6

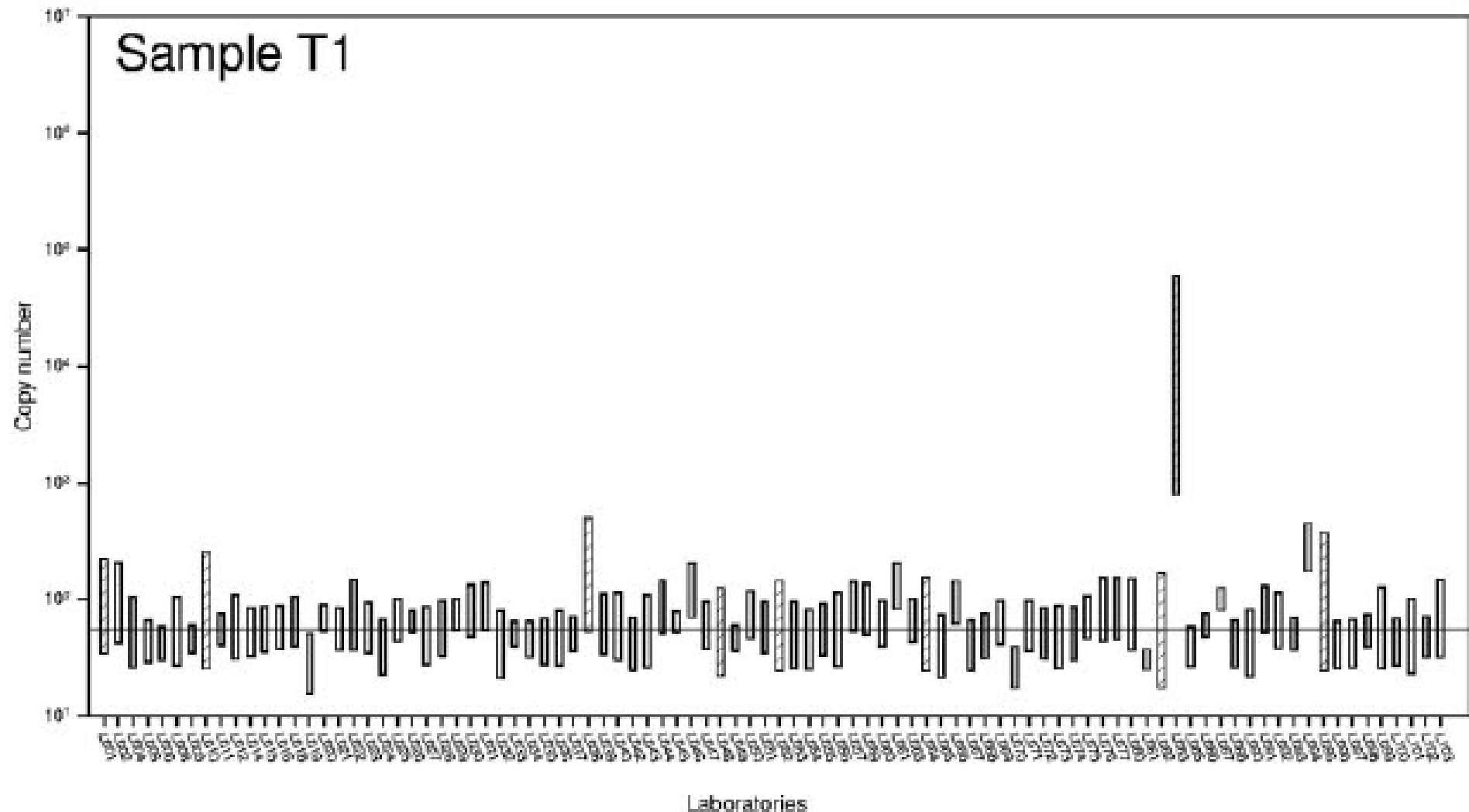


- High variability of PCR efficiency
- Differences in sizes
- Contamination issues

- EQA testing Real Time PCR
 - Used *ABL* proto oncogene
 - Participants given:
 - Primers and fluorescent probes
 - Plasmid standards (10, 10², 10³, 10⁴, 10⁵ copies/5μl)
 - 3 unknown test samples (cloned cDNAs)
 - Laboratories asked to:
 - Construct a calibration curve
 - Estimate cDNA copy numbers

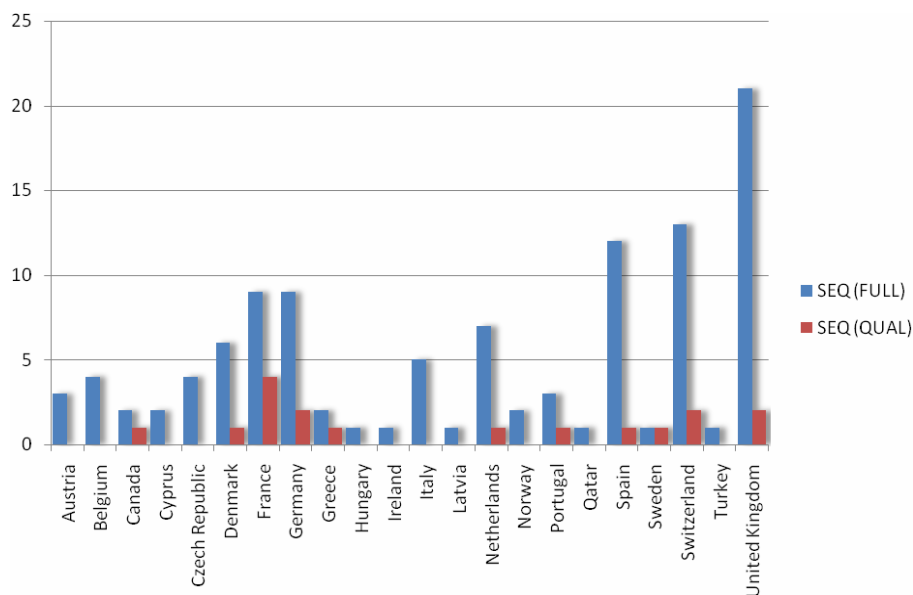
EQUAL-quant

- 93 labs returned results and 74 met performance criteria



Generic Technical EQA schemes

- DNA Sequencing
 - 6 years : 2002 - 2008



The European Molecular Genetics Quality Network

FINAL REPORT

EMQN External Quality
Assessment scheme for DNA
Sequencing (2007)

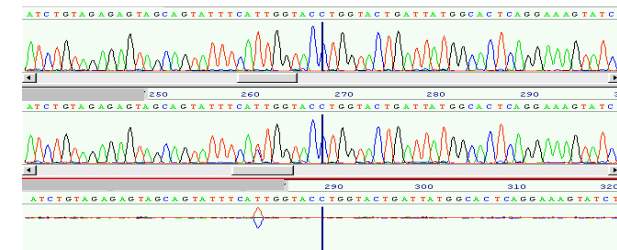
DNA Sequencing EQA scheme

- **Materials**

- Unpurified PCR products
 - CFTR, BRCA, OCRL-1 genes
 - Normal control, Normal, Heterozygote, Homozygote & a deletion Heterozygote
- Sequencing primers provided

- **Validation:**

- 2 independent laboratories



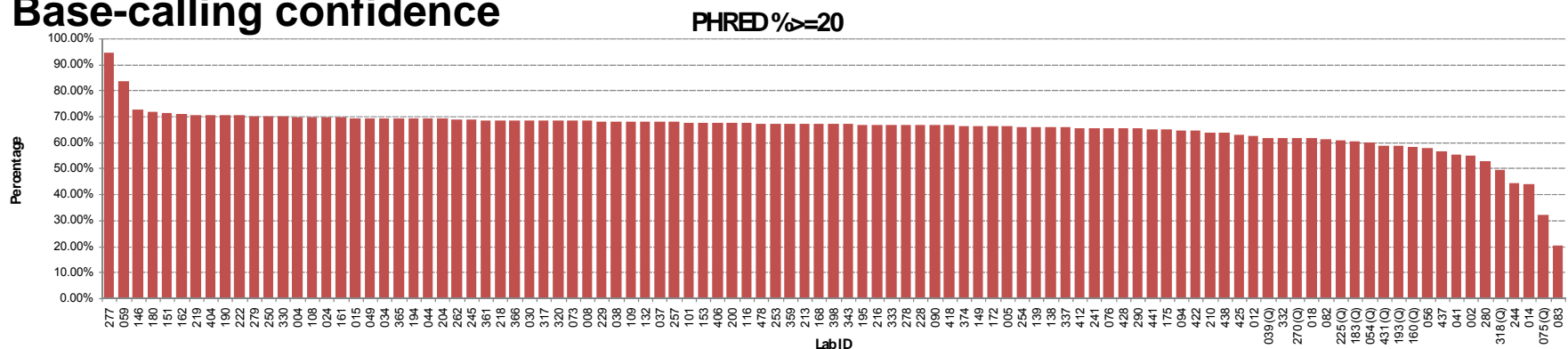
Assessment of Genotyping & Interpretation

SAMPLE	GENOTYPE	SCORE	HGVS INTERPRETATION	ACCEPTABLE VARIATIONS	REF SEQ
SEQ07_01	Mutation present c.[733G>A] + [=]	2.00	c.[733G>A] + [=]	c.[733G>A] c.733G>A	NM_000276.3
			p.[Gly245Arg] + [=]	p.[Gly245Arg] p.Gly245Arg p.[G245R] p. G245R	NM_000276.3
SEQ07_02	Mutation present c.[729dupT]	2.00	c.[729dupT]	None	NM_000276.3
			p.[Val244CysfsX13]	p.[V244CfsX13]+p.[V244CfsX13] p.[Val244CysfsX13] p.[V244CfsX13] p.[Val244fs] p.[V244fs] p.Val244fs p.V244fs	NM_000276.3
SEQ07_03	Mutation absent c.[=] + [=]	2.00	c.[=] + [=]	c.[=] c.= Wild type WT	NM_000276.3
			p.[=] + [=]	p.[=] p.=	NM_000276.3

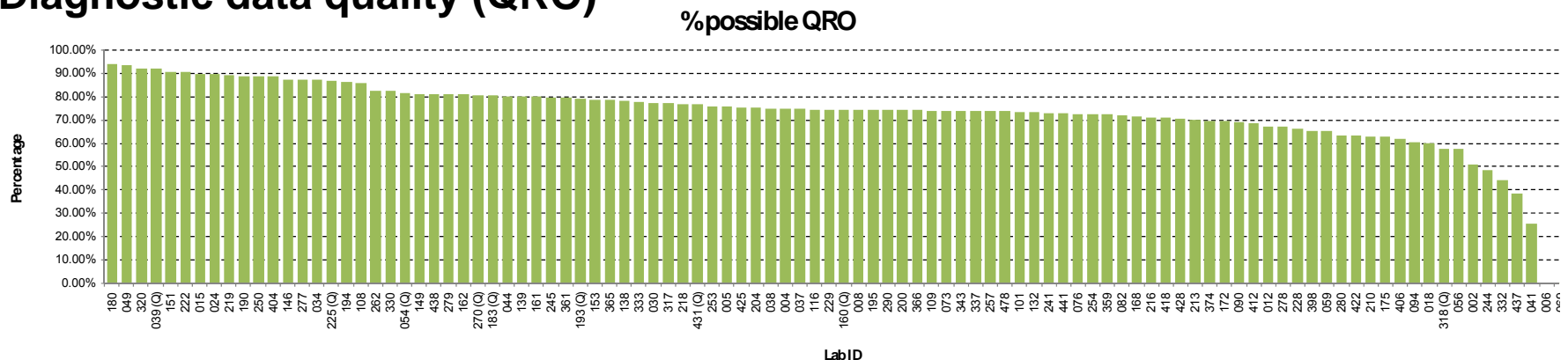
Quality of raw data

- 5 different parameters
 - PHRED scores (20,30,40), Quality Read Length (QRL), Quality Read Overlap (QRO)

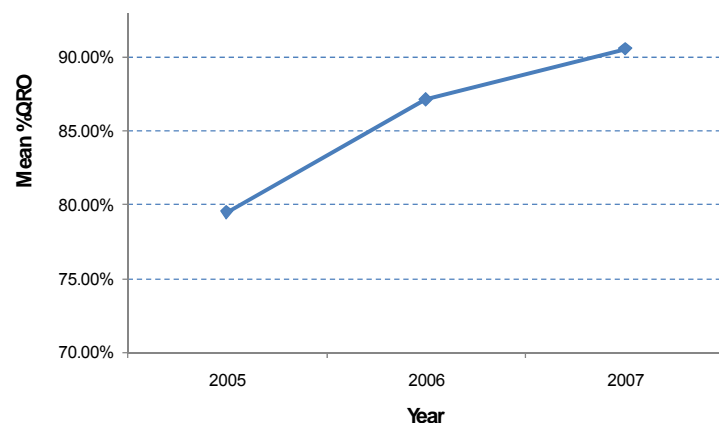
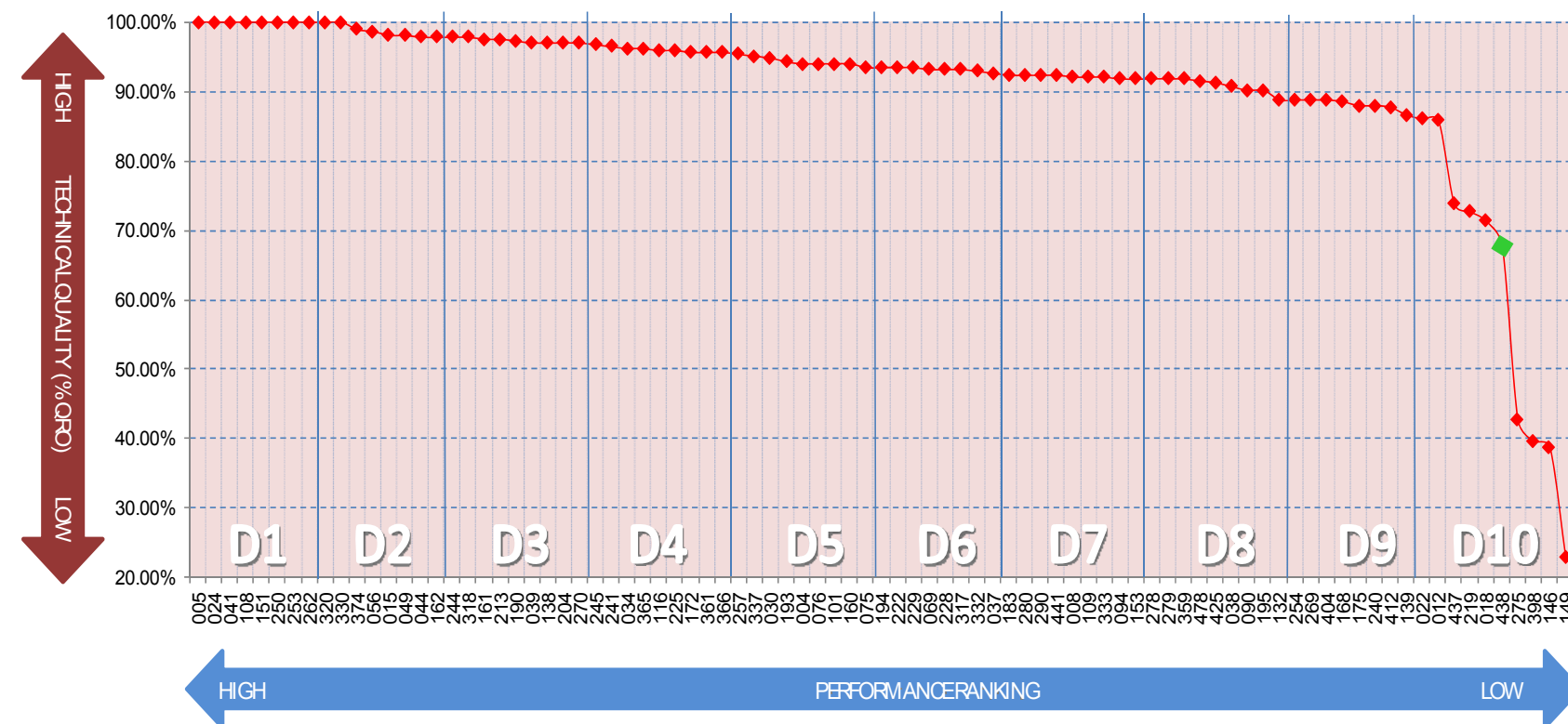
Base-calling confidence



Diagnostic data quality (QRO)



Benchmarking data quality



Benchmarking DNA sequence quality
Patton et al., *Clinical Chemistry*. 2006;52:728-736



Learning from EQA?

2007 scheme scores

Scheme	No of reports	Av. genotyping score	Av. interpretation score	No of errors leading to misdiagnosis
Total	788	1.91 (max 2.00)	1.78 (max 2.00)	45 (5.7%)
DMD	45	1.87	1.83	7
FRAX	73	1.94	1.88	3
FRDA	34	1.96	1.78	1
HFE	60	1.96	1.82	2
HNPCC	60	1.93	1.8	3
HD	55	1.94	1.81	3
SMA	34	1.94	1.9	0

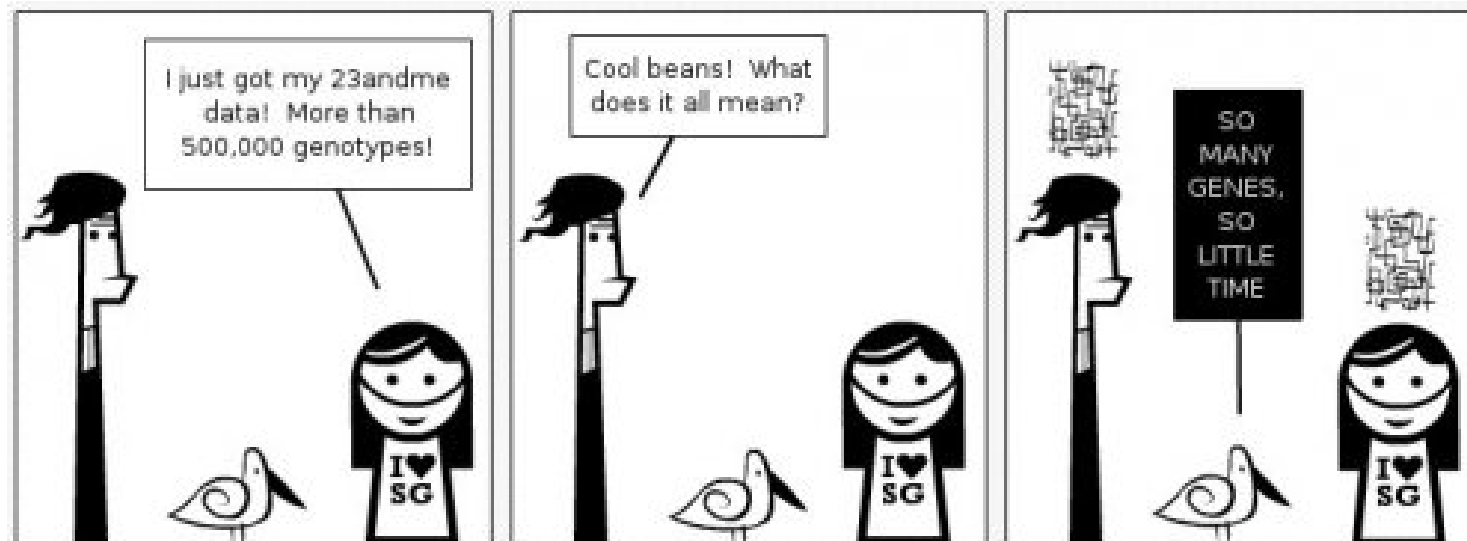
Error rate = number of genotyping errors over all returns

Common types of errors

- Wrong name / date of birth
- Sample mix up
 - 10 incidences in 2008
- Incorrect genotype
- Incorrect interpretation
- Incorrect nomenclature

Summary

- Best practice issues
 - What quality needed for reporting
 - SNP
 - Unclassified variants
- EQA
 - Single disease specific schemes
 - Issues specific for disorder (nomenclature etc)
 - Reporting and interpretation also assessed
 - Technique specific EQA
 - Errors are still being made



Acknowledgments

- Rob Elles, Simon Patton, Simon Ramsden
- EMQN management team
- EMQN scheme organisers and assessors
- Participating laboratories
- EuroGentest