npg

ARTICLE

Evaluation of PCR-based preimplantation genetic diagnosis applied to monogenic diseases: a collaborative ESHRE PGD consortium study

Jos Dreesen[1,11], Aspasia Destouni[2,11], Georgia Kourlaba[3], Birte Degn[4], Wulf Christensen Mette[4], Filipa Carvalho[5], Celine Moutou[6], Sioban Sengupta[7], Seema Dhanjal[7], Pamela Renwick[8], Steven Davies[9], Emmanouel Kanavakis[2], Gary Harton[10,12] and Joanne Traeger-Synodinos*[,2]

Preimplantation genetic diagnosis (PGD) for monogenic disorders currently involves polymerase chain reaction (PCR)-based methods, which must be robust, sensitive and highly accurate, precluding misdiagnosis. Twelve adverse misdiagnoses reported to the ESHRE PGD-Consortium are likely an underestimate. This retrospective study, involving six PGD centres, assessed the validity of PCR-based PGD through reanalysis of untransferred embryos from monogenic-PGD cycles. Data were collected on the genotype concordance at PGD and follow-up from 940 untransferred embryos, including details on the parameters of PGD cycles: category of monogenic disease, embryo morphology, embryo biopsy and genotype assay strategy. To determine the validity of PCR-based PGD, the sensitivity (Se), specificity (Sp) and diagnostic accuracy were calculated. Stratified analyses were also conducted to assess the influence of the parameters above on the validity of PCR-based PGD. The analysis of overall data showed that 93.7% of embryos had been correctly classified at the time of PGD, with Se of 99.2% and Sp of 80.9%. The stratified analyses found that diagnostic accuracy is statistically significantly higher when PGD is performed on two cells versus one cell ($P = 0.001$). Se was significantly higher when multiplex protocols versus singleplex protocols were applied ($P = 0.005$), as well as for PGD applied on cells from good compared with poor morphology embryos ($P = 0.032$). Morphology, however, did not affect diagnostic accuracy. Multiplex PCR-based methods on one cell, are as robust as those on two cells regarding false negative rate, which is the most important criteria for clinical PGD applications. Overall, this study demonstrates the validity, robustness and high diagnostic value of PCR-based PGD.
European Journal of Human Genetics (2014) 22, 1012–1018; doi:10.1038/ejhg.2013.277; published online 4 December 2013

Keywords: PCR; preimplantation genetic diagnosis; diagnostic efficiency evaluation

INTRODUCTION

Preimplantation genetic diagnosis (PGD) is a widely established reproductive alternative for couples with a high-risk of transmitting an inherited disorder. With respect to monogenic diseases, PGD can theoretically be applied for any genetic disease with a definitive molecular diagnosis and/or defined marker linkage within a family.[1,2] According to 10 years of data collection by the ESHRE PGD Consortium, PGD has been applied to over 190 different monogenic disorders.[3] Currently all methods for PGD to exclude monogenic diseases are based on the polymerase chain reaction (PCR), with or without prior whole-genome amplification. Genetic analysis may be performed at various stages post fertilization, including the oocyte/zygote biopsied on the first day post-insemination (polar body analysis), on 1–2 blastomeres from

cleavage-stage embryos biopsied on the third day post-insemination (blastomere biopsy) or on 5–10 trophectoderm cells biopsied from blastocysts on the fifth day post-insemination (blastocyst biopsy). To date, most PGD cycles have used blastomere biopsy, as in this study.

A common characteristic between all biopsy stages is the limited quantity of sample available for genetic analysis, usually a single cell only. It is this aspect of PGD that has been the most technically challenging, potentially compounded by the often sub-optimal quality of the embryo and/or embryo cell biopsied. The innate limitations of single-cell PCR include total PCR failure, allelic drop-out (ADO), and sample contamination. Overall, PCR protocols have to fulfil many conditions: they should have rapid turnaround time, be robust, sensitive and above all, absolutely accurate, to preclude misdiagnosis.

[1]Departments of Clinical Genetics and School for Oncology and Developmental Biology, GROW, Maastricht University Medical Centre, Maastricht, The Netherlands; [2]Laboratory of Medical Genetics, University of Athens, St Sophia's Children's Hospital, Athens, Greece; [3]The Stavros Niarchos Foundation-Collaborative Center for Clinical Epidemiology and Outcomes Research (CLEO), First and Second University Department of Pediatrics, National and Kapodistrian University of Athens, School of Medicine, Athens, Greece; [4]Fertility Clinic, Department of Obstetrics and Gynaecology, Aarhus University Hospital, Skejby DK-8200 Aarhus N, Denmark; [5]Department of Genetics, Faculty of Medicine, University of Porto, Porto, Portugal; [6]Université de Strasbourg FRANCE & HUS – Service de la Biologie de la Reproduction, CMCO, 19, Rue Louis Pasteur, BP120, Schiltigheim, France; [7]UCL Centre for PG&D, Institute for Women's Health, London, UK; [8]Guy's & St Thomas' Centre for Preimplantation Genetic Diagnosis and Genetics, Guy's & St Thomas' NHS Foundation Trust, London, UK; [9]Embryogenesis Assisted Conception Unit, Athens, Greece; [10]Reprogenetics, 3 Regent Street, Suite 301, Livingston, NJ, USA
*Correspondence: Dr J Traeger-Synodinos, Laboratory of Medical Genetics, University of Athens, St Sophia Children's Hospital, Athens 11527, Greece. Tel: +30 213 2013 460; Fax: +30 210 7795 553; E-mail: jtraeger@med.uoa.gr
[11]The first two authors contributed equally to the manuscript and should be regarded as joint First Authors.
[12]Current address: BlueGnome, 11781 Lee Jackson Memorial Highway, Suite 190, Fairfax, VA 22033, USA.
Received 21 March 2013; revised 16 September 2013; accepted 24 October 2013; published online 4 December 2013

Misdiagnosis can be classified as adverse or benign.[4] Adverse misdiagnosis includes cases in which there is the initiation of an affected pregnancy or the birth of an affected child. Benign misdiagnoses generally include a normal diagnosis in an embryo in fact heterozygous for a recessive disorder. In addition, healthy embryos, which are misdiagnosed as affected, can be treated as a category of adverse misdiagnosis, as this type of misdiagnosis reduces the number of embryos available for transfer, potentially decreasing the success of the IVF-PGD treatment outcome. During 10 years of data collection recording the outcome of over 4700 cycles for monogenic PGD, 12 adverse misdiagnoses were reported.[3,4] However, this is probably an underestimation, as many embryo transfers have no follow up (no pregnancy or birth). Furthermore, based on a survey done amongst ESHRE PGD consortium members in 2008 (unpublished results), only a minority of centres perform reanalysis of untransferred supernumerary embryos. Reasons for this are mainly attributed to limited staff and/or funds, no access to untransferred spare embryos (relevant for most PGD centres offering 'transport' PGD), and in some countries, legislation which forbids biopsy of the embryo (eg in Germany).[5,6]

As audit is invaluable, the primary objective of this multi-centre study was to identify the validity of PCR-based PGD by comparing results at the time of PGD with the results of the embryo follow-up analysis in a large cohort of samples. The secondary objective was to identify factors which may influence the validity of PCR-based PGD, including the embryo biology, the PCR-PGD genotyping strategies, the number of cells used in the PGD analysis, and the category of monogenic disease for which the PGD was applied.

MATERIALS AND METHODS
Study design and study population
The Embryo-Follow-up study was facilitated by the ESHRE PGD Consortium as a multi-centre study, which aimed to retrospectively evaluate the validity of PCR-based PGD. The study was conducted between October 2009 and May 2010. Specifically, an invitation was sent out to all ESHRE PGD Consortium centres which had contributed data to the annual data collections (>50). Of these, 23 initially responded and finally 6 centres submitted data that met the inclusion and data integrity criteria. The inclusion criteria for centres were defined as the submission, by the set deadline, of information on at least 50 reanalyzed embryos through the mandatory completion of all fields in the database (Table 1) in the coded format requested. The parameters in the database included general information, (centre, disease category), embryo information for the PGD cycle, genotype analysis for PGD and reanalysis, conditions of embryo reanalysis, genotype assay type, and outcome(s) (Table 1). All data was evaluated for integrity, leading to complete information on 940 reanalysed embryos.

According to regional regulations, each participating centre obtained ethical approval as well as informed consent from all couples donating embryos included in the study. All couples had undergone PGD as they were at risk of transmitting a monogenic disorder to their offspring.

Embryos considered as eligible for inclusion in this study were those that had been genotyped during a clinical PGD cycle but were not suitable or required for transfer or cryopreservation based on: (a) genetic unsuitability based on the PGD-derived genotype (affected), (b) poor developmental capacity and morphology (assessed by the embryologist), and (c) a couples' decision that their supernumerary embryos were not required for further reproductive treatment cycles. Thus, reanalysis occurred on day 4–5 post fertilization and all centres ensured randomized selection of embryos included for reanalysis to minimize bias in the study cohort. None of the embryos in this study had been previously reanalysed. To eliminate intra-centre variation, supernumerary whole embryos or cell fractions of embryos were reanalyzed using the exact same PCR-based PGD protocols applied to genotype blastomeres at PGD. Moreover, each PGD centre applied their standard assisted reproduction technology (ART) and PGD procedure according to the recommendations outlined in the ESHRE Best Practice Guidelines.[7–9]

Data analysis and statistics
Reanalysis genotypes in whole embryos or cell fractions (>2 cells) of embryos were defined as the 'true' genotype and hence 'true' embryo status, on the assumption that the analysis of multiple cells is more accurate. Based on PGD analysis and embryo reanalysis genotypes, blastomeres and reanalysed embryos were classified as affected, unaffected and aberrant. Blastomeres and embryos could be potentially scored with aberrant genotypes if the PCR protocol applied for PGD and follow-up involved linkage marker analysis (two or more markers), which demonstrated abnormal ploidy (eg monosomy, trisomy).

To determine the validity of PCR-based PGD, the sensitivity (Se), specificity (Sp) and diagnostic accuracy were calculated following the categories presented in Table 2, all with a 95% confidence interval (95% CI). The Se was defined as the proportion of affected/aberrant embryos diagnosed correctly by PGD (true positive), whereas the Sp was defined as the proportion of unaffected embryos diagnosed correctly by PGD (true negative). The diagnostic accuracy of PCR-based PGD was calculated as the proportion of embryos whose genotype results at reanalysis were in agreement with the results at PGD (ie true positive and true negative results). For the embryos whose results at PGD did not match the results at reanalysis (ie false-positive, FP, or false-negative, FN, results), the possible cause of discordance at PGD was evaluated from the genotype result and recorded. The potential causes were: (i) ADO, (ii) mosaicism, (iii) contamination, and (iv) other.

The effect of various parameters on the validity of PCR-based PGD, Se (95% CI), Sp (95% CI) and diagnostic accuracy (95% CI) were also calculated. The parameters included the impact of embryo biology, the PCR-PGD protocol strategies followed, the number of cells analysed during the PGD, and the category of disease transmission (autosomal recessive (AR), autosomal dominant (AD), X-linked dominant (XL-D) or recessive (XL-R)) for which the PGD was applied. To this end the data was divided into the following subgroups: (i) multiplex PCR method, (ii) singleplex PCR method, (iii) 1 cell biopsy, (iv) 2 cell biopsy, (v) good morphology (vi) poor morphology, (vii) AR, (viii) AD, (ix) XL-R and (x) XL-D.

The impact of embryo biology was assessed by grouping the embryos according to embryo morphology scoring. All centres scored the embryo morphology according to the same criteria[10] assigning embryos to good, intermediate and poor groups. For comparisons within the morphology subgroups, only data on the best versus the poorest morphology subgroups were compared in order to minimize overlap. Although, the participating centres used the same criteria, embryo scoring is subjective and thus could lead to overlap between intermediate embryology morphology groups.[10]

To further assess the diagnostic performance parameters of multiplex versus singleplex PCR-based PGD methods along with one versus two cell biopsy the following subgroups were created and compared: Singleplex PCR on 1 cell or 2 cell versus Multiplex PCR on 1 cell or 2 cell biopsy. Additionally, performance parameters were calculated for each centre to identify potential differences among centres.

Moreover, as genotype aberrations in blastomeres and embryos are attributed to a biological phenomenon and not a technical limitation of the PGD-PCR protocol (such as ADO or contamination), a further analysis was conducted after excluding the aberrant embryos (808 entries remained). The existence of significant differences in Se, Sp and accuracy among subgroups were examined by using the Fisher exact test. However, owing to multiple comparisons, the Bonferroni correction was used in order to account for the increase in Type I error.

Finally, simple and multiple logistic regressions were used to evaluate the association of various characteristics with the diagnostic accuracy of PGD-PCR. Because of the cluster design of the current study (more than one embryo was enrolled in the study by each centre), the centre was considered in these analyses as a cluster variable. The results are presented as odds ratio (OR) and 95% CI, and a probability value of 5% was considered as statistically significant. STAT software was used for all the statistical calculations (version 8; 2003 Corp, College Station, TX, USA).

Table 1 List of fields included in database

Field title in database	Data entry
General information	
Centre number	
Disease name	
Disease category	1 = Autosomal recessive; 2 = Autosomal dominant, 3 = X-linked recessive, 4 = X-linked dominant
Cycle number	
Embryo information for PGD cycle	
Embryo ID	Embryo identification for PGD cycle
Fertilization PN number	0, 1, 2, 3
Day PGD biopsy	
Number cells on which PGD results based	1, 2, etc
Embryo morphology[a] grade at PGD	1 = best, & 4 = poorest
Genetic analysis for PGD and reanalysis	
PGD genotype	M = pathological allele; N = normal
PGD embryo status	1 = not affected; 2 = affected; 3 = aberrant[b]
Reanalysis genotype	M = pathological allele; N = normal
Reanalysis embryo status	1 = not affected; 2 = affected; 3 = aberrant[b]
Conditions of reanalysis, assay type, and outcome(s)	
Day reanalysis	
Embryo morphology[a] grade at reanalysis	1 = best, & 4 = poorest
Number cells reanalysed	indicate if whole embryo
Reanalysis result CONCORDANT STATUS	Y/N
Possible reason discordancy	1 = ADO; 2 = contamination; 3 = mosaicism; 4 = other
Assay type	1 = multiplex; 2 = singleplex

[a]Embryo morphology was categorized in to four subgroups, with grade 1 being the quartile embryos of best morphology, and grade 4 the quartile embryos of poorest morphology (according to the Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group of Embryology consensus for embryo morphology assessment scoring criteria in each centre[10]).
[b]Aberrant: This refers to analysis using linkage marker protocols, when an embryo may be classified as likely having abnormal ploidy (eg trisomy) based on marker analysis; the genotype for the specific disease may also be 'unaffected'.

Table 2 Embryo status PGD *versus* embryo status reanalysis

	Affected/aberrant embryos at reanalysis	Unaffected embryos at reanalysis	Total
Affected/aberrant embryos at PGD	a (TP)	b (FP)	a + b
Unaffected embryos at PGD	c (FN)	d (TN)	c + d
Total	a + c	b + d	N

Abbreviations: TP, true positive; FN, false negative; TN, true negative; FP, false positive.
Sensitivity: TP/(TP + FN).
Specificity: TN/(TN + FP).
Diagnostic accuracy: TP + TN/(TP + FN + TN + FP).

RESULTS

The study sample included results from 940 reanalysed PGD cases for 53 different genetic disorders (15 AR, 24 AD, 10 XL-R and 4 XL-D.

Data submitted by the participants showed that 56.8% of the PCR-based PGD protocols were performed on one biopsied cell, 72.9% of the embryo genotypes were achieved by application of a multiplex PCR-based protocol and that 25.2% of the embryos were assigned to the 'good morphology' category at PGD.

Overall diagnostic performance parameters of PCR-PGD

Among the 940 reanalysed embryos, results for the genetic status at the time of PGD showed that 234 blastomere-based embryo analysis were classified as unaffected, 590 as affected and 116 as aberrant. Embryo reanalysis showed that 283 embryos were unaffected, 578 were affected and 79 aberrant. Hence, in 881 (out of 940) reanalyzed embryos, the status at the time of PGD was concordant with the reanalysis embryo status (diagnostic accuracy of 93.7%). The overall Se and Sp of the PCR-PGD methods were 99.2 and 80.9%, respectively (Table 3). Fifty-nine embryos were misclassified at PGD (5 were classified as FN and 54 as FP). With regards to the cause of discordant results, the observed five FN entries were attributed to mosaicism. The majority of FP results were attributed to mosaicism (29), followed by ADO (17), contamination (7), and 'other causes' (1).

Subgroup data analysis

Stratified analysis by centre revealed no significant difference among centres in terms of Se ($P = 0.700$, Table 3). However, statistically significant differences were detected in Sp and diagnostic accuracy ($P = 0.002$ and $P < 0.001$, respectively). In particular, Sp was found to range between 69% and 94%, with the exception of one centre where only three unaffected embryos were included, all of which were FP (Sp: 0%). In terms of diagnostic accuracy, a range of 84–98% was detected across all participating centres.

Moreover, stratified analysis of the PCR-PGD protocols applied (multiplex *versus* singleplex) showed that multiplex protocols perform statistically significantly better than singleplex protocols in terms of Se (99.8 *versus* 97.9%, $P = 0.03$), whereas no significant difference was detected in Sp and diagnostic accuracy ($P = 0.352$ and $P = 0.547$, respectively) (Table 3). Concerning the number of biopsied cells that underwent PCR-based PGD, the analysis showed that two-cell biopsy exhibits a significant advantage in terms of diagnostic accuracy compared with one-cell biopsy (96.7 *versus* 91.6%, $P = 0.001$).

Table 3 Validity of PCR-PGD analysis compared with embryo reanalysis (*n* = 940 embryos, including aberrant)

	No. of subjects				Validity		
	TP	FN	TN	FP	Se (95% CI)	Sp (95% CI)	Accuracy (95% CI)
Overall	652	5	229	54	99.2% (98.2–99.8%)	80.9% (75.8–85.3%)	93.7% (91.9–95.2%)
By centre							
E	42	0	15	1	100% (91.6–100%)	93.8% (69.8–99.8%)	98.3% (90.8–99.9%)
C	46	0	0	3	100% (92.3–100%)	–	93.9% (83.1–98.7%)
A	355	5	98	29	98.6% (96.8–99.5%)	77.2% (68.9–84.1%)[a]	93.0% (90.4–95.1%)
D	21	0	22	8	100% (83.9–100%)	73.3% (54.1–87.7%)[a]	84.3% (71.4–93.0%)[a]
G	154	0	76	5	100% (97.6–100%)	93.8% (86.1–98.0%)	97.9% (95.1–99.3%)
B	34	0	18	8	100% (89.7–100%)	69.2% (48.2–85.7%)[a]	86.7% (75.4–94.0%)[a]
No. of cells used in PGD							
One cell	337	3	152	42	99.1% (97.4–99.8%)	78.3% (71.9–83.9%)	91.6% (88.8–93.8%)[b]
Two cells	300	2	76	11	99.3% (97.6–99.9%)	87.4% (78.5–93.5%)	96.7% (94.3–98.2%)[b]
>2 cells (3 or 4)	15	0	1	1	100% (78.2–100%)	50% (1.2–98.7%)	94.1% (71.3–99.9%)
Disease category							
AR	123	1	96	26	99.2% (95.6–99.9%)	78.7% (70.4–85.6%)	89.0% (84.4–92.6%)[c]
AD	442	3	104	25	99.3% (98.0–99.9%)	80.6% (72.7–87.0%)	95.1% (93.0–96.7%)[c]
XL-R	43	1	20	2	97.8% (88.0–99.9%)	90.9% (70.8–98.9%)	95.4% (87.3–99.0%)
XL-D	44	0	9	1	100% (91.9–100%)	90.0% (55.5–99.7%)	98.1% (90.1–99.9%)
PCR-PGD protocol							
Singleplex	192	4	45	14	97.9% (94.8–99.4%)[b]	76.3% (63.4–86.4%)	92.9% (89.1–95.8%)
Multiplex	460	1	184	40	99.8% (98.8–99.9%)	82.1% (76.5–86.9%)	94.0% (92.0–95.7%)
Combined biopsy–PGD-PCR strategies							
Singleplex one cell	67	3	21	8	95.7% (87.9–99.1%)[c]	72.4% (52.8–87.3%)	88.9% (81.0–94.3%)[c]
Singleplex two cells	118	1	24	5	99.2% (95.4–99.9%)	82.8% (64.2–94.2%)	95.9% (91.4–98.5%)
Multiplex one cell	270	0	131	34	100% (98.6–100%)[c]	79.4% (72.4–85.3%)	92.1% (89.2–94.5%)
Multiplex two cells	182	1	52	6	99.5% (97.0–100%)	89.6% (78.8–96.1%)	97.1% (94.1–98.8%)[c]
Embryo morphology							
Class 1	178	0	44	15	100% (97.9–100%)[b]	74.6% (61.6–85.0%)	93.7% (89.8–96.4%)
Class 4	37	2	21	1	94.9% (82.7–99.4%)	95.5% (77.1–99.9%)	95.0% (86.3–98.9%)

[a]*P* < 0.05 for comparison with the centre G after Bonferroni correction.
[b]*P* < 0.05.
[c]*P* < 0.05 after Bonferroni correction.

Seventeen embryos whose PGD was based on three or four cells were excluded from the statistical analysis (only descriptive results are presented) due to the extremely small sample size.

The combined diagnostic efficiency of PCR-based PGD strategies (molecular method and biopsy protocol) was investigated by comparing the following subgroups: Singleplex 1 cell (S1cell), Singleplex2 cells (S2cell), Multiplex 1 cell (M1cell), and Multiplex 2 cells (M2cell) biopsy. A statistically significant difference was observed between the S1 cell and M1 cell for the Se (*P* = 0.048), whereas there was no significant difference detected for Se in the remaining pairwise comparisons (Table 3). In terms of diagnostic accuracy, multiplex PGD with two cells seems to identify the status of embryos with significantly greater accuracy compared with singleplex PGD with one cell (97.1 *versus* 88.9%, *P* = 0.024), whereas a marginally higher diagnostic accuracy was detected in multiplex PGD with two cells compared with multiplex PGD with one cell (97.1 *versus* 92.1%, *P* = 0.066).

To investigate the degree of the influence of embryo morphology of PCR-based PGD tests on the diagnostic outcomes, data from embryos

from the two extreme categories: good (class 1), and poor (class 4) morphology (see 'Methods' section) were compared. The Se of PCR-based PGD seems to be statistically significantly higher among embryos with 'good' morphology (100%) compared with those with 'poor' morphology (94.9%, *P* = 0.032). On the other hand, no significant effect of embryo morphology on Sp and diagnostic accuracy of PCR-based PGD was detected (*P* = 0.057 and *P* = 0.999, respectively) (Table 3).

Finally, no statistically significant difference was demonstrated in Se and Sp among the mode of inheritance subgroups, although Sp was significantly lower in AR subgroup (78.7%) compared with XL-D subgroup (90.0%). However, in terms of diagnostic accuracy, it was found to be statistically significantly higher in AD group compared with AR group (89.0 *versus* 95.1%, *P* = 0.002) (Table 3).

Multiple logistic regression, after taking into account the potential clustering effect of the centre, revealed that the probability of agreement between PGD and reanalysis results (either true positive or true negative) is statistically significantly higher when two cells as opposed to one cell are used in PGD (*P* < 0.001), when multiplex

1016

instead of singleplex protocols are applied ($P = 0.001$) and when the type of disease for which the PGD is applied is AD and XL-D *versus* AR ($P < 0.001$) (Table 4).

Exclusion of all aberrant embryos

The analysis of the data following exclusion of all aberrant genotype results demonstrated that 785 out of the 808 blastomeres, genotyped at PGD had concordant embryo status with the corresponding reanalyzed embryos, giving Se, Sp, and diagnostic accuracy of 99.6, 91.6, and 97.2%, respectively (Table 5). The FN group contained two samples, which were scored discordant due to mosaicism. In the FP group (21 samples), a variety of phenomena underlie discordant diagnostic outcomes. ADO (9/21) was the most frequent reason of discrepancy between PGD and reanalysis embryo status and under-lying genotypes. Other causes of FP results were mosaicism (7/21) and contamination (4/21), whereas 1 could not be explained by the above.

The results of stratified analyses are also presented in Table 5. These results indicate that excluding aberrant embryos, the only parameter, which was found to have statistically significant difference was the diagnostic accuracy between AR and AD ($P = 0.006$, Table 5).

Multiple logistic regression, confirmed the absence of statistically significant differences, for all subgroups (see Materials and Methods) with the exception of the disease mode of inheritance for which the PGD was applied, whereby a statistically significant difference in the probability of agreement between PGD and reanalysis results was detected between AR and XL-D ($P < 0.001$, Table 6).

Table 4 Factors affecting the diagnostic accuracy of PGD; Results from simple and multiple logistic regression by using centre as a cluster factor ($N = 923^a$)

	Crude OR (95% CI)	Adjusted OR (95% CI)
No. of cells used in PGD		
One cell	Ref	Ref
Two cells	**2.66 (1.35–5.24)**	**2.64 (1.71–4.06)**
Disease category		
AR	Ref	Ref
AD	**2.40 (1.77–3.26)**	**2.51 (1.50–4.20)**
XL-R	2.58 (0.87–7.69)	2.48 (0.85–7.26)
XL-D	**6.53 (2.90–14.72)**	**7.92 (4.54–13.80)**
PCR-PGD protocol		
Multiplex	Ref	Ref
Singleplex	0.84 (0.52–1.35)	**0.50 (0.27–0.92)**
Embryo morphology		
Class 1	Ref	Not included[b]
Class 4	1.30 (0.30–5.74)	
Combined biopsy-PGD-PCR strategies		
Singleplex one cell	Ref	Not included[c]
Singleplex two cells	**2.96 (1.77–4.94)**	
Multiplex one cell	1.47 (0.63–3.41)	
Multiplex two cells	**4.18 (2.20–7.90)**	

Bold reflects statistically significant results.
[a]Embryos, which underwent multiple biopsy at PGD (>2 cell biopsy) were excluded from the current analysis.
[b]Morphology of embryos was not included in the multiple logistic regression because this variable refers to a subsample and its inclusion in the analysis would reduce significantly the sample size.
[c]This variable was not included in the multiple logistic regression as the number of cells and the type of PCR-PGD protocol have been included in the model as two independent variables.

DISCUSSION

The primary objective of the present study was the investigation of the diagnostic accuracy and validity of PCR-based PGD protocols that are routinely applied to diagnose single gene diseases. It is the first multicentre study to collectively quantify diagnostic performance parameters, based on data analysis methodology previously used in a single centre to internally audit the accuracy and validity of PCR-based PGD protocols.[11] In general, the results are expected to highlight pitfalls of PCR-based PGD (technical or biological), which can then be addressed to optimize clinical PGD results.

Overall, data analysis demonstrates the validity, robustness, and high diagnostic performance of the PCR-based PGD protocols. The Se and Sp and accuracy of PCR-based PGD protocols applied to diagnose single gene disorders are high (99.2, 80.9, and 93.7%). The high Se reflects the significantly low risk of adverse misdiagnosis. This observation is very important as adverse misdiagnosis may have severe consequences upon the couples, such as the initiation and subsequent termination of an affected pregnancy or the birth of an affected child.[4] The inter-centre comparison was not shown to influence the Se of PGD results. However, Sp was significantly different amongst the participating centres owing to the differences in the FP rate. This could be attributed to the general trend to overestimate affected status when interpreting the PCR-based PGD results, in order to preclude transfer of any affected embryos.

The calculated FN rate stems from five FN diagnostic outcomes, which were all attributed to mosaicism. This provides a strong indication that the protocols used to date are efficiently designed to detect the genetic change that causes the disease but that adverse misdiagnosis could stem from testing a cell, which does not represent the genetic make-up of the entire embryo. It has been previously shown that almost 50% of the embryos created by ARTs are mosaic, containing both normal (diploid) and abnormal (non-diploid) cells.[12–15] With respect to FP results, the contribution of mosaicism to the FP rate ($FP = 19.1\%$) in the overall data, reaches 54%. FP diagnosis is characterized as more benign as it does not lead to the initiation of an affected pregnancy, although it may decrease the number of genetically transferable embryos, hence reducing the probability of achieving a pregnancy. Data showed that the second most frequent cause of FP diagnosis was ADO, an inherent pitfall of single-cell PCR. Strategies to overcome the technical limitations of single-cell PCR have been previously described[4,16–18] and incorporated in the ESHRE best practice guidelines for monogenic PGD.[7] To preclude adverse misdiagnosis due to ADO and contamination, the co-amplification of linked polymorphic markers across the locus of interest is highly recommended and the advantages of the approach had been extensively described.[19,20] In addition, the analogous effectiveness of multiple displacement amplification methods in preimplantation genetic haplotyping has been reported.[20,21] The technical superiority of multiplex over singleplex approaches is definitively demonstrated by the results of this present study.

With respect to biopsy strategies, the analysis of two biopsied cells from a single embryo showed higher OR compared with one-cell biopsy. However, the Se and hence the FN rate for both strategies remains the same. In addition, two-cell biopsy strategies appear to be more efficient than one-cell biopsy in terms of lower FP diagnosis. Overall, multiplex protocols on one cell (M1cell) present a signifi-cantly better diagnostic approach than singleplex one cell (S1cell). As could be expected, the latter demonstrated the poorest diagnostic performance parameters. S1cell methods had the highest FN and FP rates, thus having the greatest risk of transferring an affected embryo

Table 5 Validity of PCR-PGD analysis compared with embryo reanalysis ($n = 808$ embryos, excluding aberrant)

	No. of subjects				Validity		
	TP	FN	TN	FP	Se (95% CI)	Sp (95% CI)	Accuracy (95% CI)
Overall	556	2	229	21	99.6% (98.7–99.9%)	91.6% (87.4–94.7%)	97.2% (95.8–98.2%)
By centre							
E	42	0	15	1	100% (91.6–100%)	93.8% (69.7–99.8%)	98.3% (90.8–99.9%)
C	37	0	0	0	100% (90.5–100%)	–	100% (90.5–100%)
A	282	2	98	8	99.3% (97.5–99.9%)	92.5% (85.7–96.7%)	97.4% (95.3–98.8%)
D	21	0	22	6	100% (83.9–100%)	78.6% (59.0–91.7%)	87.8% (75.2–95.4%)
G	154	0	76	5	100% (97.6–100%)	93.8% (86.2–98.0%)	97.8% (95.1–99.3%)
B	20	0	18	1	100% (83.2–100%)	94.7% (74.0–99.9%)	97.4% (86.5–99.9%)
No. of cells used in PGD							
One cell	282	0	152	15	100% (98.7–100%)	91.0% (85.6–94.9%)	96.7% (94.5–98.1%)
Two cells	261	2	76	5	99.2% (97.3–99.9%)	93.8% (86.2–98.0%)	98.0% (95.8–99.2%)
>2 cells (3 or 4)	13	0	1	1	100% (75.3–100%)	50.0% (1.2–98.7%)	93.3% (68.0–99.8%)
Disease category							
AR	97	1	96	12	98.9% (94.4–99.9%)	88.9% (81.4–94.1%)	93.7% (89.4–96.6%)[a]
AD	382	0	104	7	100% (99.0–100%)	93.7% (87.4–97.4%)	98.6% (97.1–99.4%)[a]
XL-R	41	1	20	1	97.6% (87.4–99.9%)	95.2% (76.2–99.9%)	96.8% (89.0–99.9%)
XL-D	36	0	9	1	100% (90.2–100%)	90.0% (55.5–99.7%)	97.8% (88.5–99.9%)
PCR-PGD protocol							
Singleplex	158	1	45	6	99.4% (96.5–99.9%)	88.2% (76.1–95.6%)	96.7% (93.2–98.6%)
Multiplex	398	1	184	15	99.7% (98.6–99.9%)	92.5% (87.9–95.7%)	97.3% (95.7–98.5%)
Combined biopsy-PGD-PCR strategies							
Singleplex one cell	52	0	21	3	100% (93.2–100%)	87.5% (67.6–97.3%)	96.1% (88.9–99.2%)
Singleplex two cells	100	1	24	2	99.0% (94.6–99.9%)	92.3% (74.9–99.0%)	97.6% (93.3–99.5%)
Multiplex one cell	230	0	131	12	100% (98.4–100%)	91.6% (85.8–95.6%)	96.8% (94.4–98.3%)
Multiplex two cells	161	1	52	3	99.4% (96.6–99.9%)	94.5% (84.9–98.8%)	98.2% (95.3–99.5%)
Embryo morphology							
Class 1	146	0	46	2	100% (97.5–100%)	95.8% (85.7–99.5%)	99.0% (96.3–99.9%)
Class 4	36	1	21	1	97.3% (85.8–99.9%)	95.5% (77.2–99.9%)	96.6% (88.3–99.6%)

[a]$P < 0.05$ after Bonferroni correction.

or significantly decreasing the number of embryos available for transfer (50% of FPs caused by ADO).

However, it must be emphasized that multiplex PCR-based methods applied on 1 cell (M1cell) are as robust as those on two cells (M2cell) regarding FN results, which is of fundamental importance for clinical PGD. The higher FP rate when analysing a single *versus* two cells are likely due to the inability in the former to double-check results and resolve ambiguous genotyping results on a second cell. Although the FP rate is higher in the M1cell subgroup, in practice a balance has to be found between genotype accuracy and good embryo integrity. It has been reported that two-cell biopsy can be detrimental to both the embryo development and clinical outcome,[22,23] and as the ultimate desired outcome of a PGD cycle is the initiation and delivery of an unaffected pregnancy, M1cell protocols fulfil this balance. Two-cell biopsy should probably be considered when a reliable multiplex PCR-based method is hard to develop, such as for PGD cases with *de novo* mutations when informative markers cannot be found.[7]

The present study also investigated whether embryo morphology influences PCR-based PGD diagnostic validity. Although, the Se between good and poor morphology embryos was found to be

statistically different, further statistical analysis (simple and multiple logistic regression) showed that morphology does not affect the diagnostic accuracy of PGD. Morphology subgroup comparisons should be treated with caution due to the small size of data analysed.

Finally, this study quantified the impact of embryos scored with aberrant genotypes on the diagnostic efficiency of PCR-based PGD protocols. In general, a significant difference between aberrant *versus* non-aberrant embryos was exhibited between the Sp (<0.001) and accuracy (<0.001). FP entries, attributed to mosaicism and ADO, were decreased dramatically through the removal of aberrant embryos. More specifically, when aberrant embryos were included the main causes of FP genotypes were attributed to mosaicism (29) or ADO (17), whereas after their exclusion FP genotypes were accounted for by 7 cases of mosaicism and 9 ADO. On the other hand, no significant differences were demonstrated in the Se ($P = 0.375$) between these two data sets, indicating that, despite of the presence of aberrant embryos, current PGD protocols are well suited to detect embryos affected for a monogenic disease.

Furthermore, the exclusion of aberrant embryos led to the abolition of significant differences between subgroups, with the exception of the accuracy between AD *versus* AR mode of inheritance subgroups

Table 6 Factors affecting the diagnostic accuracy of PGD: results from simple and multiple logistic regression by using centre as a cluster factor ($N = 792^a$)

	Crude OR (95% CI)	Adjusted OR (95% CI)
No. of cells used in PGD		
One cell	Ref	Ref
Two cells	1.66 (0.56–4.91)	1.16 (0.46–2.92)
Disease category		
AR	Ref	Ref
AD	**4.67 (1.59–13.71)**	6.31 (0.98–40.87)
XL-R	2.05 (0.59–7.10)	2.09 (0.54–8.14)
XL-D	**3.03 (1.04–8.87)**	**3.90 (2.45–6.20)**
PCR-PGD protocol		
Multiplex	Ref	Ref
Singleplex	0.80 (0.32–1.97)	0.50 (0.13–1.94)
Embryo morphology		
Class 1	Ref	
Class 4	0.29 (0.04–2.23)	
Combined biopsy-PGD-PCR strategies		
Singleplex one cell	Ref	Not included[b]
Singleplex two cells	1.69 (0.94–2.06)	
Multiplex one cell	1.23 (0.44–3.46)	
Multiplex two cells	2.19 (0.93–5.13)	

Bold reflects statistically significant results.
[a]The aberrant embryos and those whose PGD result was based on more than 2 cells were excluded from the current analysis.
[b]This variable was not included in the multiple logistic regression as the number of cells and the type of PCR-PGD protocol have been included in the model as two independent variables.

whereby the FP rate was higher in the AR subgroup. In the case of PGD for an AR disease, in which one pathological allele is definitively identified but there is no result for the trans allele (either due to biological or technical reasons), there is only a 50% chance that the embryo is unaffected, and thus the tendency is to score the embryo as genetically affected. On the other hand, when an AD disease is tested, if the only allele detected is the mutant allele, then the embryo will be correctly scored as affected.

The present study is the first multicentre evaluation of the degree of clinical validity of PCR-based PGD methods. The statistical analysis took into consideration limitations associated with the study, such as the embryo morphology scoring which is relatively subjective, and accepts that the response rate among the centres invited to participate relative to the number of centres who participated ($\sim 11\%$) was slightly lower than expected.

In conclusion, this study demonstrates the validity, robustness, and high diagnostic value and performance of a wide range of PCR-based methods currently used in clinical PGD. In addition, it provides substantial evidence that the embryo biology has a significant impact on the diagnostic accuracy of PCR-based PGD methods and this should always be taken into account when designing the strategy of methodologies and evaluating genotype results for monogenic PGD.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

1 SenGupta SB, Delhanty JD: Preimplantation genetic diagnosis: recent triumphs and remaining challenges. *Expert Rev Mol Diagn* 2012; **12**: 585–592.
2 Fiorentino F: Molecular genetic analysis of single cells. *Semin Reprod Med* 2012; **30**: 267–282.
3 Harper JC, Wilton L, Traeger-Synodinos J et al: The ESHRE PGD Consortium: 10 years of data collection. *Hum Reprod Update* 2012; **18**: 234–247.
4 Wilton L, Thornhill A, Traeger-Synodinos J, Sermon KD, Harper JC: The causes of misdiagnosis and adverse outcomes in PGD. *Hum Reprod* 2009; **24**: 1221–1228.
5 Augst C: Regulating dangerous futures: the German Embryo Protection Act of 1990–legislation in risk society. *Soc Leg Stud* 2000; **9**: 205–226.
6 Benohr-Laqueur S: Fighting in the legal grey area: an analysis of the German Federal Court of Justice decision in case preimplantation genetic diagnosis. *Poiesis Prax* 2011; **8**: 3–8.
7 Harton GL, De Rycke M, Fiorentino F et al: ESHRE PGD consortium best practice guidelines for amplification-based PGD. *Hum Reprod* 2011; **26**: 33–40.
8 Harton GL, Harper JC, Coonen E, Pehlivan T, Vesela K, Wilton L: ESHRE PGD consortium best practice guidelines for fluorescence in situ hybridization-based PGD. *Hum Reprod* 2011; **26**: 25–32.
9 Harton GL, Magli MC, Lundin K, Montag M, Lemmen J, Harper JC: ESHRE PGD Consortium/Embryology Special Interest Group – best practice guidelines for polar body and embryo biopsy for preimplantation genetic diagnosis/screening (PGD/PGS). *Hum Reprod* 2011; **26**: 41–46.
10 Alpha Scientists in Reproductive, M., E.S.I.G.o. Embryology. The Istanbul consensus workshop on embryo assessment: proceedings of an expert meeting. *Hum Reprod* 2011; **26**: 1270–1283.
11 Dreesen J, Drusedau M, Smeets H et al: Validation of preimplantation genetic diagnosis by PCR analysis: genotype comparison of the blastomere and corresponding embryo, implications for clinical practice. *Mol Hum Reprod* 2008; **14**: 573–579.
12 Bielanska M, Jin S, Bernier M, Tan SL, Ao A: Diploid-aneuploid mosaicism in human embryos cultured to the blastocyst stage. *Fertil Steril* 2005; **84**: 336–342.
13 Bielanska M, Tan SL, Ao A: Chromosomal mosaicism throughout human preimplantation development in vitro: incidence, type, and relevance to embryo outcome. *Hum Reprod* 2002; **17**: 413–419.
14 van Echten-Arends J, Mastenbroek S, Sikkema-Raddatz B et al: Chromosomal mosaicism in human preimplantation embryos: a systematic review. *Hum Reprod Update* 2011; **17**: 620–627.
15 Mertzanidou A, Wilton L, Cheng J et al: Microarray analysis reveals abnormal chromosomal complements in over 70% of 14 normally developing human embryos. *Hum Reprod* 2012; **28**: 256–264.
16 Fiorentino F, Biricik A, Nuccitelli A et al: Strategies and clinical outcome of 250 cycles of preimplantation genetic diagnosis for single gene disorders. *Hum Reprod* 2006; **21**: 670–684.
17 Altarescu G, Brooks B, Kaplan Y et al: Single-sperm analysis for haplotype construction of de-novo paternal mutations: application to PGD for neurofibromatosis type 1. *Hum Reprod* 2006; **21**: 2047–2051.
18 Altarescu G, Eldar-Geva T, Varshower I et al: Real-time reverse linkage using polar body analysis for preimplantation genetic diagnosis in female carriers of de novo mutations. *Hum Reprod* 2009; **24**: 3225–3229.
19 Renwick P, Ogilvie CM: Preimplantation genetic diagnosis for monogenic diseases: overview and emerging issues. *Expert Rev Mol Diagn* 2007; **7**: 33–43.
20 Renwick P, Trussler J, Lashwood A, Braude P, Ogilvie CM: Preimplantation genetic haplotyping: 127 diagnostic cycles demonstrating a robust, efficient alternative to direct mutation testing on single cells. *Reprod Biomed Online* 2010; **20**: 470–476.
21 Renwick PJ, Trussler J, Ostad-Saffari E et al: Proof of principle and first cases using preimplantation genetic haplotyping – a paradigm shift for embryo diagnosis. *Reprod Biomed Online* 2006; **13**: 110–119.
22 De Vos A, Staessen C, De Rycke M et al: Impact of cleavage-stage embryo biopsy in view of PGD on human blastocyst implantation: a prospective cohort of single embryo transfers. *Hum Reprod* 2009; **24**: 2988–2996.
23 Haapaniemi Kouru K, Malmgren H, Nordenskjold M, Fridstrom M, Csemiczky G, Blennow E: One-cell biopsy significantly improves the outcome of preimplantation genetic diagnosis (PGD) treatment: retrospective analysis of 569 PGD cycles at the Stockholm PGD centre. *Hum Reprod* 2012; **27**: 2843–2849.