Received: 2 November 2022 | Revised: 22 December 2022 | Accepted: 31 January 2023

DOI: 10.1002/pd.6333

ORIGINAL ARTICLE

PRENATAL DIAGNOSIS WILEY

Non-invasive fetal genotyping for maternal alleles with droplet digital PCR: A comparative study of analytical approaches

Joe Shaw[1] | Elizabeth Scotchman[1] | Ben Paternoster[1] | Maureen Ramos[1] | Sarah Nesbitt[1] | Sophie Sheppard[1] | Tristan Snowsill[2] | Lyn S. Chitty[1,3] | Natalie Chandler[1]

[1]North Thames Genomic Laboratory Hub, Great Ormond Street NHS Foundation Trust, London, UK

[2]Health Economics Group, University of Exeter, Exeter, UK

[3]Genetic and Genomic Medicine, UCL Great Ormond Street Institute of Child Health, London, UK

Correspondence
Lyn S. Chitty
Email: l.chitty@ucl.ac.uk

Funding information
NIHR Biomedical Research Centre at Great Ormond Street Hospital; National Institute for Health and Care Research

Abstract

Objectives: To develop a flexible droplet digital PCR (ddPCR) workflow to perform non-invasive prenatal diagnosis via relative mutation dosage (RMD) for maternal pathogenic variants with a range of inheritance patterns, and to compare the accuracy of multiple analytical approaches.

Methods: Cell free DNA (cfDNA) was tested from 124 archived maternal plasma samples: 88 cases for sickle cell disease and 36 for rare Mendelian conditions. Three analytical methods were compared: sequential probability ratio testing (SPRT), Bayesian and z-score analyses.

Results: The SPRT, Bayesian and z-score analyses performed similarly well with correct prediction rates of 96%, 97% and 98%, respectively. However, there were high rates of inconclusive results for each cohort, particularly for z-score analysis which was 31% overall. Two samples were incorrectly classified by all three analytical methods; a false negative result predicted for a fetus affected with sickle cell disease and a false positive result predicting the presence of an X-linked *IDS* variant in an unaffected fetus.

Conclusions: ddPCR can be applied to RMD for diverse conditions and inheritance patterns, but all methods carry a small risk of erroneous results. Further evaluation is required both to reduce the rate of inconclusive results and explore discordant results in more detail.

Key points

What is already know about this topic?

- Digital PCR can be used to determine the fetal genotype of single gene disorders by dosage analysis of cfDNA in maternal plasma.
- Previous studies have shown the feasibility of this approach with small sample cohorts but reported rates of incorrect results are as high as 20%.

- Different statistical analysis methods have been reported, but their comparative accuracy has not been assessed.

What does this study add?

- We report the results for digital PCR analysis on over 120 cfDNA samples and a large-scale comparison of three different analytical approaches.
- Accounting for technical variation using z-score analysis can significantly reduce incorrect fetal genotype predictions, but all three methods tested have a small residual error rate.

1 | INTRODUCTION

During pregnancy, cell free fetal DNA (cffDNA) is shed from the placenta and can be detected in maternal plasma alongside cell free DNA (cfDNA) from maternal tissues.[1] This has allowed safer access to fetal genetic material and the development of prenatal testing using only a maternal blood sample, including screening for common aneuploidies,[2] determination of fetal sex[3] and fetal *RHD* status,[4] and non-invasive prenatal diagnosis (NIPD) of single gene disorders.[5]

Unlike detection of de novo or paternally inherited variants, which are not present in the maternal genome, determining the fetal inheritance of maternal variants is technically challenging, as the variant of interest is also present in the maternal cfDNA. This has complicated the development of NIPD for scenarios in which the mother is affected with a dominant condition or is a carrier for an X-linked or recessive condition. Detecting fetal inheritance of maternal alleles can be achieved using relative haplotype dosage analysis (RHDO), in which next generation sequencing (NGS) is used to determine the inheritance of parental haplotypes in the cfDNA.[6] However, the requirement of DNA from the biological father and a familial proband currently limits its availability.

Notably, RHDO is not available for sickle cell disease (SCD), despite this being a life-limiting autosomal recessive (AR) condition, with over 300,000 patients born globally every year[7] and one of the most frequent requests for prenatal diagnosis in the United Kingdom. Approximately 70% of SCD patients are homozygous for a single missense variant in the *HBB* gene (NM_000518.5:c.20A>T), referred to as the haemoglobin S (HbS) allele.[8] Each year, over 300 invasive prenatal tests are performed for SCD in the United Kingdom alone, and a sample from the biological father is not received in up to 40% of cases when requested[9]; this would prevent access to RHDO even if it was available.

Pregnant women who are heterozygous for rare disease variants are also not served by RHDO, as it is too expensive to permit validation for rare genetic conditions in a publicly funded healthcare system. Moreover, due to the requirement of a high number of informative single nucleotide polymorphisms (SNPs) to differentiate haplotypes at the gene locus, RHDO is currently not offered to consanguineous couples.[10] There is therefore an unmet clinical need for a cost-effective and flexible non-invasive method for detecting maternally inherited variants using only a maternal blood sample. Patients welcome the prospect of such tests but stress that accuracy is of paramount importance for any new method to be an acceptable alternative to invasive testing.[11]

Fetal inheritance of maternal alleles can also be detected using relative mutation dosage (RMD), which measures imbalances in the abundance of variant and normal sequences in the cfDNA.[12] Detecting changes in allelic dosage due to the fetal genotype is complicated by the low concentration of total cfDNA, which is present between 600 and 2000 genome equivalents per millilitre of plasma (GE/ml) in early pregnancy.[13,14] Of this total cfDNA, the fetal DNA may comprise less than 10%, and any allelic dosage change will be relative to this fetal fraction. For a cfDNA test to be a clinical alternative to invasive testing, it should be available at an equivalent time-point to chorionic villous sampling (CVS), which is 11–14 weeks gestation, when the fetal fraction and absolute concentration of cfDNA are both low.[14]

Digital PCR (dPCR) has previously been applied to RMD approaches, including cohorts of haemophilia,[15,16] monogenic diabetes[17] and inherited deafness.[18] However, Barrett et al.[19] reported in a large cohort study for SCD that 7 out of 59 fetal genotype predictions using sequential probability ratio testing (SPRT) were incorrect, including one false positive and four false negatives. Recently, Sawakwongpra et al.[20] reported a misclassification rate of roughly 20% (5/24) when SPRT was applied to non-invasive fetal genotyping of beta-thalassemia. The high rate of these incorrect predictions and uncertainty about their aetiology have so far hindered the clinical implementation of RMD in cfDNA testing. Many publications have applied SPRT to RMD,[12,15,16,19,21–23] whilst z-score methods[24,25] and Bayesian approaches,[18,26] including Markov chain Monte-Carlo (MCMC) analysis,[17] have also been reported. However, cohort sizes are often limited, and no large-scale comparison of analytical approaches has been performed.

In this study, we developed a flexible droplet digital PCR (ddPCR) workflow to perform NIPD for maternal pathogenic variants with X-linked, autosomal dominant (AD) and autosomal recessive (AR) inheritance, including the common *HBB* c.20A>T variant, and a case from a consanguineous family. We then applied three analytical methods, SPRT, Bayesian and z-score analyses, and compared their accuracy in predicting fetal genotypes.

2 | METHODS

2.1 | Sample collection

Plasma samples were collected from women, and biological fathers if available, attending fetal medicine units between 2012 and 2021, as

part of the RAPID project (NIHR RP-PG-0707-10107; Research Ethics Committee reference: 14/SC/1020).[13] All samples were pseudo-anonymised and consent for research obtained. For this study, 127 samples were identified with the following criteria: samples from singleton pregnancies from women who were heterozygous for pathogenic variants in Mendelian disease genes and with a fetal genotype determined by invasive prenatal or postnatal testing. Three cases were subsequently excluded due to evidence of the haemoglobin C (HbC) allele, sample contamination and the presence of a twin pregnancy which was not noted on the referral. Thus 124 samples were subsequently analyzed: 88 for SCD (*HBB* c.20A>T) and 36 for a range of variants in different disease genes (Figure 1).

Maternal plasma samples were processed as previously described[13] and stored at −80°C. cfDNA was extracted using a QiaSymphony instrument using the QIAsymphony DSP Circulating Nucleic Acid Kit (Qiagen). Genomic DNA (gDNA) for maternal and paternal controls was extracted from stored blood pellets using a QuickGene-610L (Kurabo) kit. Further details are included in Supplementary Methods.

2.2 | Droplet digital PCR

Variant discrimination assays were designed for the 35 pathogenic variants present in our cohort: 20 for X-linked conditions, 6 for AD conditions and 9 for AR conditions, including the common *HBB* c.20A>T variant for SCD. Primer and probe sequences for the *FGFR3* c.1138G>A,[27] *CFTR* c.1521_1523del[28] and *HBB* c.126_129del[12] variants were taken from previous publications. The sequences, amplicons and annealing temperatures for each assay are shown in Supplementary Table 1, and PCR conditions are included in the Supplementary Methods.

ddPCR was performed on a Bio-Rad QX200 system with an automated droplet generator (Bio-Rad). Each assay was optimised on gDNA from heterozygous parental controls using an annealing temperature gradient. Each cfDNA sample was then tested with the ddPCR assay for the relevant pathogenic variant, and results were analyzed using Quantasoft (v1.7.4). The fetal fraction was determined via ddPCR for the *ZFY* locus[29] or a paternally inherited SNP.[21,30] Identification of informative SNPs was performed using NGS and ddPCR and is described in more detail in the Supplementary Methods. Parental gDNA was tested alongside the cfDNA at equivalent concentrations to provide a comparison dataset of samples known to be truly heterozygous.

2.3 | Limit of detection experiment

A limit of detection experiment was designed to test the sensitivity of the *HBB* c.20A>T ddPCR assay. gDNA from patients confirmed by Sanger sequencing to be homozygous for either the HbS (HbSS) or HbA (HbAA) alleles was fragmented to an average size of 150 bp on a Covaris E220 Ultrasonicator, and the DNA fragment profile was assessed using an Agilent 2200 TapeStation. HbAA and HbSS gDNA was spiked into a HbAS gDNA sample at increments from 2% to 12%, at different total concentrations of DNA. These mixtures simulated the composition of cfDNA from a pregnant woman bearing affected and unaffected fetuses with varying fetal fractions and cfDNA concentrations. This experiment was performed only for the SCD assay due to sample availability.

2.4 | Analysis

SPRT was performed as previously described for autosomal[31] and X-linked inheritance patterns,[15] using a likelihood ratio threshold of 8.[32] The Bayesian analysis was performed for AD variants as described by Caswell et al.,[17] with additional models for X-linked and AR inheritance, and a threshold of 0.95[17] for fetal genotype classification. Finally, we performed a z-score analysis, which was modified from Chiu et al.[2] Heterozygous parental gDNA results were used as a control dataset, with known equal concentrations of variant and reference alleles. The z-score was then calculated as the number of standard deviations by which the cfDNA sample variant fraction differed from the mean of the heterozygous gDNA controls. Applying the same thresholds as Chiu et al.,[2] z-scores greater than 3 or less than −3 were used to predict homozygous and hemizygous fetal genotypes, whilst z-scores between 2 and −2 were predicted to have heterozygous fetal genotypes.

3 | RESULTS

Overall, 124 samples were analyzed from 116 different families, for 35 pathogenic variants in 28 genes (Figure 2). Samples were collected from between 9 and 35 weeks gestation (median: 12 + 4) with fetal fractions ranging from 1.8% to 31.6% (median: 6.6%) (Supplementary Table 2).

3.1 | Assay optimisation

The limit of detection study using sonicated gDNA showed that the SCD assay could distinguish 4% spike-ins of both HbSS and HbAA at DNA inputs ranging from 3000 to 12000 molecules (Supplementary Figure 1). We then looked at the variant fraction of the heterozygous gDNA controls with known equal concentrations of variant and reference alleles. These showed substantial variation (43.8%–55.4%), particularly when fewer than 2000 haploid genome equivalents (GE) were measured, due to the sampling error associated with low DNA inputs. The variation was such that when the SPRT and Bayesian analyses were applied to the control dataset, several replicates exceeded the classification thresholds (Supplementary Figure 2).

Based on the results of this gDNA testing and the limit of detection experiment for the SCD assay, we set additional quality filters across both cohorts: samples with fewer than 2000 GE at the

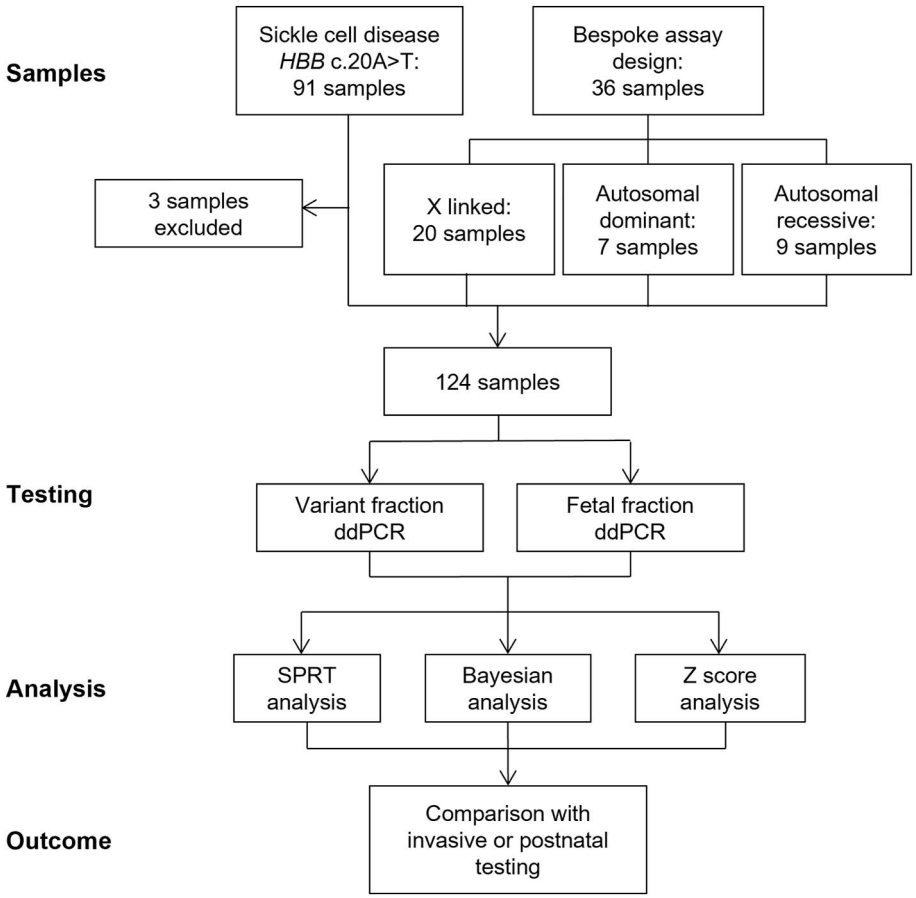

FIGURE 1 The testing workflow of the sample cohort. The numbers in each box indicate the number of samples tested. Three samples included for sickle cell disease testing were subsequently excluded due to the detection of contamination, the presence of the haemoglobin C allele and due to a twin pregnancy.

variant site or with a fetal fraction below 4% were classified as inconclusive for all three analytical methods. The 4% fetal fraction threshold is consistent with that routinely used for non-invasive prenatal screening for aneuploidies[33,34] and in our laboratory for complex cfDNA assays for monogenic conditions. This was therefore applied across all the assays, not just the SCD assay, rather than testing bespoke assays due to lack of sample availability. A total of 25 samples, 13 in the SCD cohort and 12 in the bespoke cohort, did not meet these quality criteria and were therefore classified as inconclusive across all three analytical methods.

3.2 | Sickle cell disease cohort

The *HBB* c.20A>T ddPCR assay was successfully optimised and additional testing assessed the impact of the common HbC allele (NM_000518.5:c.19G>A) on probe binding, which generated a distinct low fluorescence droplet cluster (Supplementary Figure 3).

When compared to the results of invasive testing, the SPRT, Bayesian and z-score analyses generated 97%, 98% and 99% correct fetal genotype predictions for reportable cases, with 22%, 25% and 23% inconclusive results, respectively (Table 1). These inconclusive rates include 13 cases, 15% of the SCD cohort, which did not meet the quality criteria due to low fetal fraction or low DNA input. There were two incorrect fetal genotype predictions generated by SPRT; a homozygous normal genotype (HbAA) predicted for a confirmed heterozygous carrier fetus (HbAS) and a heterozygous carrier genotype predicted for a homozygous affected fetus (HbSS). The Bayesian and z-score analyses generated only one of these incorrect fetal genotypes. This sample (cfDNA-122), collected at 12 + 5 weeks gestation, had a variant fraction of 50% and a fetal fraction of 6% and was incorrectly predicted to have a heterozygous fetus by all three analysis methods, whilst the CVS reported a fetus affected with SCD.

3.3 | Bespoke cohort

For the bespoke design cohort, all three analysis methods again performed similarly well, with 92% correct fetal genotype predictions for reportable cases for SPRT, 91% for the Bayesian analysis and 94% for the z-score analysis (Table 2). Among the correct predictions made by all three methods were those for the common *CFTR* c.1521_1523del variant and two common beta-thalassemia variants: *HBB* c.126_129del and *HBB* c.93-21G>A. Six samples were tested

TABLE 1 Sickle cell disease results ($n = 88$).

Fetal genotype	HbSS	HbAS	HbAA	Total	Total (%)	Percent reportable
SPRT						
Correct	12	43	12	67	76	97
Incorrect	1	1	0	2	2	3
Inconclusive	0	14	5	19	22	-
Bayesian						
Correct	11	42	12	65	74	98
Incorrect	1	0	0	1	1	2
Inconclusive	1	16	5	22	25	-
Z-score						
Correct	10	45	12	67	76	99
Incorrect	1	0	0	1	1	1
Inconclusive	2	13	5	20	23	-

Note: Results for the prediction of fetal genotypes for pregnancies at risk of sickle cell disease by sequential probability ratio testing (SPRT), Bayesian and z score analyses compared with genotypes determined from invasive or postnatal testing. Inconclusive cases are classified as not reportable.

from couples in which both parents carried the same rare variant in a recessive gene, and correct fetal genotypes were predicted by all three methods for five of these (Table 3). The overall rates of inconclusive results, including those which did not meet the quality thresholds, were slightly higher for this cohort; 33% for SPRT, 39% for the Bayesian analysis and 50% for z-score analysis. Twelve cases, 33% of the total cohort, were classified as inconclusive across all three methods due to low fetal fraction or low DNA input. Of note, seven X-linked samples had fewer than 2000 molecules detected by the variant fraction assay (cfDNA-19, cfDNA-20, cfDNA-22, cfDNA-25, cfDNA-26, cfDNA-27 and cfDNA-30) and therefore did not pass the quality filter. These samples were extracted from small plasma volumes and tested early in the study, prior to the optimisation of the cfDNA extraction method (Supporting Information S1).

The SPRT and Bayesian analyses generated the same two incorrect fetal genotype predictions for two X-linked recessive variants in the *ABCD1* and *IDS* genes; one false negative result (cfDNA-17) and one false positive result (cfDNA-29) (Table 4). The latter was also incorrectly classified by z-score analysis. This false positive result, called by all three analytical methods, was from a cfDNA sample taken at 13 + 6 weeks gestation from a woman who was heterozygous for the *IDS* c.182_189del variant, which causes mucopolysaccharidosis type 2. The *ZFX/ZFY* ddPCR assay measured a fetal fraction of 10.3%, and the variant fraction was 54.6%. The SPRT, Bayesian and z-score analyses all predicted that the fetus was hemizygous for the variant and was therefore affected. However, Sanger sequencing on CVS DNA at the time of sampling did not detect the pathogenic variant, and the fetus was born unaffected with mucopolysaccharidosis type 2.

Overall, across both the SCD and bespoke cohorts, SPRT, Bayesian and z-score analyses correctly classified 96%, 97% and 98% of reportable cases, with 25%, 29% and 30% inconclusive results, respectively (Table 5). Again, these inconclusive results include the 25 cases, 20% of the total cohort, which did not meet the quality thresholds. SPRT generated the highest number of incorrect results, with four in total across both cohorts, while z-score generated only two. A full detailed results table can be found in Supplementary Table 2.

TABLE 2 Bespoke cohort results ($n = 36$).

	Total	Total (%)	Percent reportable
SPRT			
Correct	22	61	92
Incorrect	2	6	8
Inconclusive	12	33	-
Bayesian			
Correct	20	55	91
Incorrect	2	6	9
Inconclusive	14	39	-
Z-score			
Correct	17	47	94
Incorrect	1	3	6
Inconclusive	18	50	-

Note: Results for the prediction of fetal genotypes for pregnancies at risk of rare single gene conditions by sequential probability ratio testing (SPRT), Bayesian and z score analyses compared with genotypes determined from invasive or postnatal testing. Inconclusive cases are classified as not reportable.

3.4 | X-chromosome inactivation

We hypothesised that X-chromosome inactivation (XCI) could have played a role in the false positive *IDS* variant result due to the

TABLE 3 Bespoke cohort individual results.

Inheritance	Sample number	Condition	Gene	DNA	Fetal genotype	SPRT	Bayesian	Z score
AR	cfDNA-1	Severe combined immunodeficiency	*ADA*	c.556G>A	Het	Het	Het	Het
AR	cfDNA-2	Aicardi-Goutières syndrome type 6	*ADAR*	c.2997G>T	Het	Het	Het	Het
AR	cfDNA-3	Cystic fibrosis	*CFTR*	c.1521_1523del	Hom var	Hom var	Hom var	Hom var
AR	cfDNA-4	Beta thalassemia	*HBB*	c.126_129del	Het	Het	Het	Het
AR	cfDNA-5	Beta thalassemia	*HBB*	c.93-21G>A	Het	Het	Het	Het
AR	cfDNA-6	Vitamin B12-responsive methylmalonic aciduria	*MMAA*	c.733+1G>A	Hom ref	Hom ref	Hom ref	Hom ref
AR	cfDNA-7	Congenital disorder of glycosylation type 1a	*PMM2*	c.691G>A	Hom var	Hom var	Hom var	Hom var
AR	cfDNA-8	Congenital disorder of glycosylation type 1a	*PMM2*	c.691G>A	Het	Inconclusive	Inconclusive	Inconclusive
AR	cfDNA-9	Aicardi-Goutières syndrome type 3	*RNASEH2C*	c.205C>T	Het	Het	Het	Het
AD	cfDNA-10	Hypophosphatasia	*ALPL*	c.331G>A	Het	Het	Het	Inconclusive
AD	cfDNA-11	Crouzon syndrome	*FGFR2*	c.1024T>A	Hom ref	Hom ref	Inconclusive	Hom ref
AD	cfDNA-12	Achondroplasia	*FGFR3*	c.1138G>A	Hom ref	Inconclusive	Inconclusive	Inconclusive
AD	cfDNA-13	Neurofibromatosis, type 1	*NF1*	c.6792C>A	Hom ref	Inconclusive	Inconclusive	Inconclusive
AD	cfDNA-14	Neurofibromatosis, type 1	*NF1*	c.6792C>A	Het	Het	Het	Het
AD	cfDNA-15	Treacher Collins syndrome 1	*TCOF1*	c.3611C>A	Hom ref	Hom ref	Hom ref	Hom ref
AD	cfDNA-16	von Hippel-Lindau syndrome	*VHL*	c.583C>T	Hom ref	Hom ref	Hom ref	Hom ref
XLR	cfDNA-17	Adrenoleukodystrophy	*ABCD1*	c.3G>A	Hemi var	Hemi ref	Hemi ref	Inconclusive
XLR	cfDNA-18	Menkes disease	*ATP7A*	c.1949G>A	Hemi var	Hemi var	Hemi var	Hemi var
XLR	cfDNA-19	Menkes disease	*ATP7A*	c.2916+1G>A	Hemi var	Inconclusive	Inconclusive	Inconclusive
XLR	cfDNA-20	X-linked choroideremia	*CHM*	c.1342C>T	Hemi ref	Inconclusive	Inconclusive	Inconclusive
XLR	cfDNA-21	Haemophilia A	*F8*	c.1409C>T	Hemi var	Hemi var	Hemi var	Hemi var
XLR	cfDNA-22	Haemophilia A	*F8*	c.1595G>A	Hemi var	Inconclusive	Inconclusive	Inconclusive
XLR	cfDNA-23	Haemophilia A	*F8*	c.5999G>T	Hemi var	Hemi var	Hemi var	Hemi var
XLR	cfDNA-24	Haemophilia A	*F8*	c.6046C>T	Hemi var	Hemi var	Hemi var	Hemi var
XLR	cfDNA-25	Haemophilia A	*F8*	c.6544C>T	Hemi var	Inconclusive	Inconclusive	Inconclusive
XLR	cfDNA-26	Haemophilia A	*F8*	c.6686T>C	Hemi var	Inconclusive	Inconclusive	Inconclusive
XLR	cfDNA-27	Frontometaphyseal dysplasia 1	*FLNA*	c.5169T>G	Hemi var	Inconclusive	Inconclusive	Inconclusive
XLR	cfDNA-28	IPEX syndrome	*FOXP3*	c.1010G>A	Hemi ref	Hemi ref	Hemi ref	Hemi ref
XLR	cfDNA-29	Mucopolysaccharidosis type 2	*IDS*	c.182_189del	Hemi ref	Hemi var	Hemi var	Hemi var
XLR	cfDNA-30	Mucopolysaccharidosis type 2	*IDS*	c.879+1G>T	Hemi var	Inconclusive	Inconclusive	Inconclusive
XLR	cfDNA-31	X-linked hydrocephalus	*L1CAM*	c.23T>A	Hemi var	Hemi var	Inconclusive	Inconclusive
XLR	cfDNA-32	Bartter syndrome type 5	*MAGED2*	c.1426C>T	Hemi var	Hemi var	Hemi var	Inconclusive
XLR	cfDNA-33	Norrie disease	*NDP*	c.163T>C	Hemi ref	Inconclusive	Inconclusive	Inconclusive
XLR	cfDNA-34	Ornithine transcarbamylase deficiency	*OTC*	c.905A>G	Hemi ref	Hemi ref	Hemi ref	Inconclusive
XLR	cfDNA-35	X-linked lymphoproliferative syndrome type 2	*XIAP*	c.356_359del	Hemi var	Inconclusive	Inconclusive	Inconclusive
XLD	cfDNA-36	Alport syndrome type 1	*COL4A5*	c.1295G>A	Hemi var	Hemi var	Hemi var	Inconclusive

Note: Results of droplet digital PCR for the bespoke design cohort, stratified by analysis method, compared to the results from invasive or post-natal testing. Het: heterozygous; hom var: homozygous variant; hom ref: homozygous reference; hemi ref: hemizygous reference; hemi var: hemizygous variant. AR: autosomal recessive. AD: autosomal dominant. XLR: X-linked recessive. XLD: X-linked dominant.

TABLE 4 Incorrect results.

Sample number	Variant	Fetal fraction (%)	Variant fraction (%)	Molecules detected	Fetal genotype	SPRT	Bayesian	Z score
cfDNA-17	*ABCD1* c.3G>A	7.7	49.5	3538	Hemi var	Hemi ref	Hemi ref	Inconclusive
cfDNA-29	*IDS* c.182_189del	10.3	54.6	5949	Hemi ref	Hemi var	Hemi var	Hemi var
cfDNA-87	*HBB* c.20A>T	5.2	48.5	12181	Het	Hom ref	Inconclusive	Inconclusive
cfDNA-122	*HBB* c.20A>T	6	50	8889	Hom var	Het	Het	Het

Note: Details of the four cell free DNA (cfDNA) samples with incorrect fetal genotype predictions. Molecules detected indicates the total molecules detected by the variant fraction droplet digital PCR assay for each case. Het: heterozygous; hom var: homozygous variant; hom ref: homozygous reference; hemi ref: hemizygous reference; hemi var: hemizygous variant.

TABLE 5 Total cohort results ($n = 124$).

Fetal genotype	Total	Total (%)	Percent reportable (%)
SPRT			
Correct	89	72	96
Incorrect	4	3	4
Inconclusive	31	25	-
Bayesian			
Correct	85	69	97
Incorrect	3	2	3
Inconclusive	36	29	-
Z-score			
Correct	84	68	98
Incorrect	2	2	2
Inconclusive	38	30	-

Note: Results for the prediction of fetal genotypes for all 124 pregnancies included in this study by sequential probability ratio testing (SPRT), Bayesian and z score analyses. Inconclusive results are classified as not reportable.

previously reported relationship between methylation status and cfDNA fragmentation.[35] XCI testing was performed on maternal leukocyte DNA and DNA from the CVS. The polymorphic Androgen Receptor (*AR*) (CAG)n repeat was amplified with and without HpaII digestion, demonstrating an inactivation ratio of 75:25 between the two alleles (Figure 3). The X chromosome with 25% methylation had been inherited by the unaffected male fetus. This indicated that the X chromosome bearing the normal *IDS* allele (Xn) was more commonly active in the heterozygous mother's cells than the X chromosome bearing the *IDS* c.182_189del pathogenic variant (Xv).

4 | DISCUSSION

We have developed a ddPCR workflow for RMD and applied it to a large cohort of 88 clinical samples from pregnancies at risk of SCD due to the common *HBB* c.20A>T variant, and 36 cases at risk of 26 rare genetic conditions. The SPRT and Bayesian approaches correctly classified 96% and 97% fetal genotypes in cases where a prediction was made, respectively, across both cohorts. Both methods generated a low number of incorrect fetal genotype predictions, four for the SPRT and three for the Bayesian analysis, and incorrectly classified ddPCR data from heterozygous gDNA controls. We subsequently performed a z-score analysis, which differed from the SPRT and Bayesian approaches as it compared the cfDNA result to the results of truly heterozygous parental gDNA samples. This therefore accounted for the technical variation observed in the gDNA controls and had the highest correct prediction rate for the entire cohort, 98% of genotype predictions, generating only two incorrect results.

All three analytical methods had high rates of inconclusive results, the lowest being SPRT at 25% and the highest being z-score analysis at 30%. Twenty-five samples in total, 20% of the cases, did not meet the quality criteria due to low DNA input or low fetal fraction and these were classified as inconclusive across all three analytical methods. By removing these cases and looking only at those which passed the quality criteria, the inconclusive rates are reduced to 6%, 11% and 13% for SPRT, Bayesian and z-score analyses, respectively. Since this study used archived samples, we were unable to request repeat samples at later gestations for those cases with inconclusive ddPCR results due to low fetal fraction or low numbers of molecules. Although taking a second sample later in gestation does not always resolve issues of low fetal fraction, it is expected that in a clinical scenario with access to repeat sampling in cases which fail the quality criteria, all three analysis methods may have a lower inconclusive rate. Of note, there are current clinical NIPD services being offered in the NHS reported to have high inconclusive rates during validation, which were subsequently resolved following repeat testing with later samples.[10,36]

The SPRT was developed by Wald[37] and has been applied to dosage-based cfDNA testing of sequence variants with ddPCR[12,15,16,19–21] and NGS,[22,23,38] as well as for chromosomal aneuploidy.[31] The SPRT has also been successfully applied in RHDO, where its usage is justified by the sequential nature of SNPs along the locus of interest.[6] However, as ddPCR data for a single loci is not acquired sequentially, the application of SPRT in ddPCR has been criticised[39] and incorrect fetal genotype predictions have been reported.[12,19,20,23,38] In this study, applying the SPRT to ddPCR results from heterozygous parental gDNA still generated fetal genotype

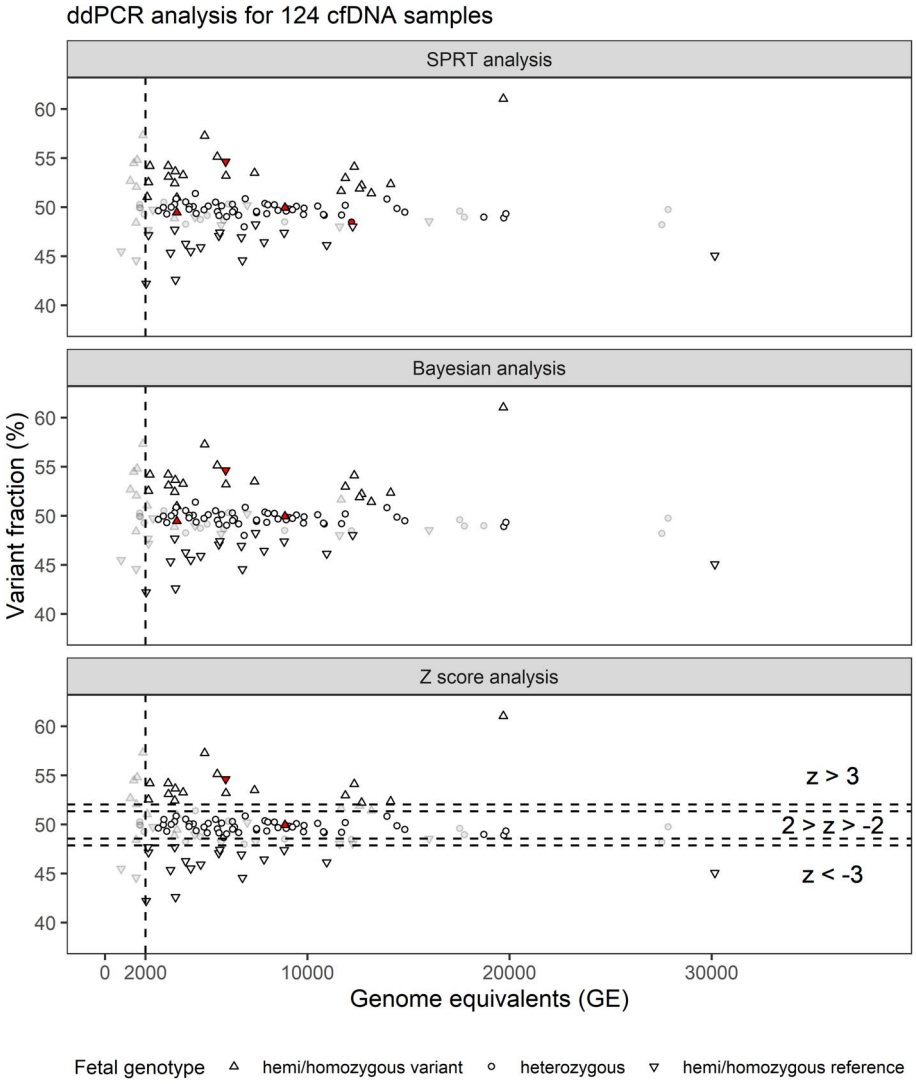

FIGURE 2 Droplet digital PCR (ddPCR) analysis for 124 cell free DNA (cfDNA) samples. Shapes indicate whether invasive testing determined the fetal genotype as hemi/homozygous for the variant (upward triangles), heterozygous (circles) or hemi/homozygous for the reference allele (downward triangles) and are colored depending on whether the prediction was correct (white), incorrect (red) or inconclusive (grey). A threshold of 2000 genome equivalents was also applied due to the high variability observed in heterozygous genomic DNA (gDNA) controls at lower quantities. For the z-score analysis, the horizontal lines show the z-score thresholds of 3, 2, −2 and −3. 95% Poisson confidence intervals are not included to prevent overplotting. [Colour figure can be viewed at wileyonlinelibrary.com]

classifications, despite no fetal signal being present. This indicates that the SPRT in this form is overly sensitive and can misinterpret technical variation as a dosage imbalance caused by cffDNA.

By contrast, the Bayesian analysis described here was developed specifically for non-invasive fetal genotyping with ddPCR by Caswell et al.[17] In a cohort of 38 pregnancies, the Bayesian approach had a correct prediction rate of 100% for fetal inheritance of maternal variants in the *GCK* or *HNF4A* genes, even at low fetal fractions. The lower accuracy of this method in our cohort can be partially attributed to the earlier gestational ages at which samples were taken, as both maternal and fetal cfDNA are lower in concentration earlier in gestation.[14] The median gestational age at sampling reported by Caswell et al. was 28 + 0 weeks gestation, whilst for our cohort it was 12 + 4 weeks.

Z-score analysis is commonly applied in non-invasive prenatal testing (NIPT) for aneuploidy, in which the dosage of a potentially trisomic chromosome, such as chromosome 21, is compared to the dosage of disomic chromosomes from the same sample.[2] We aimed to adapt this principle for single gene disorders by comparing the dosage of each pathogenic variant in the cfDNA samples with the dosage in heterozygous parental controls, and dosage results from a limit of detection study. This was the most conservative method, making slightly fewer predictions overall; however, reducing the number of incorrect results compared to the other two methods used.

Despite this reduction in incorrect fetal genotype predictions using z-score analysis, two samples, one in each cohort, were still incorrectly classified by all applied methods, and we are unable to

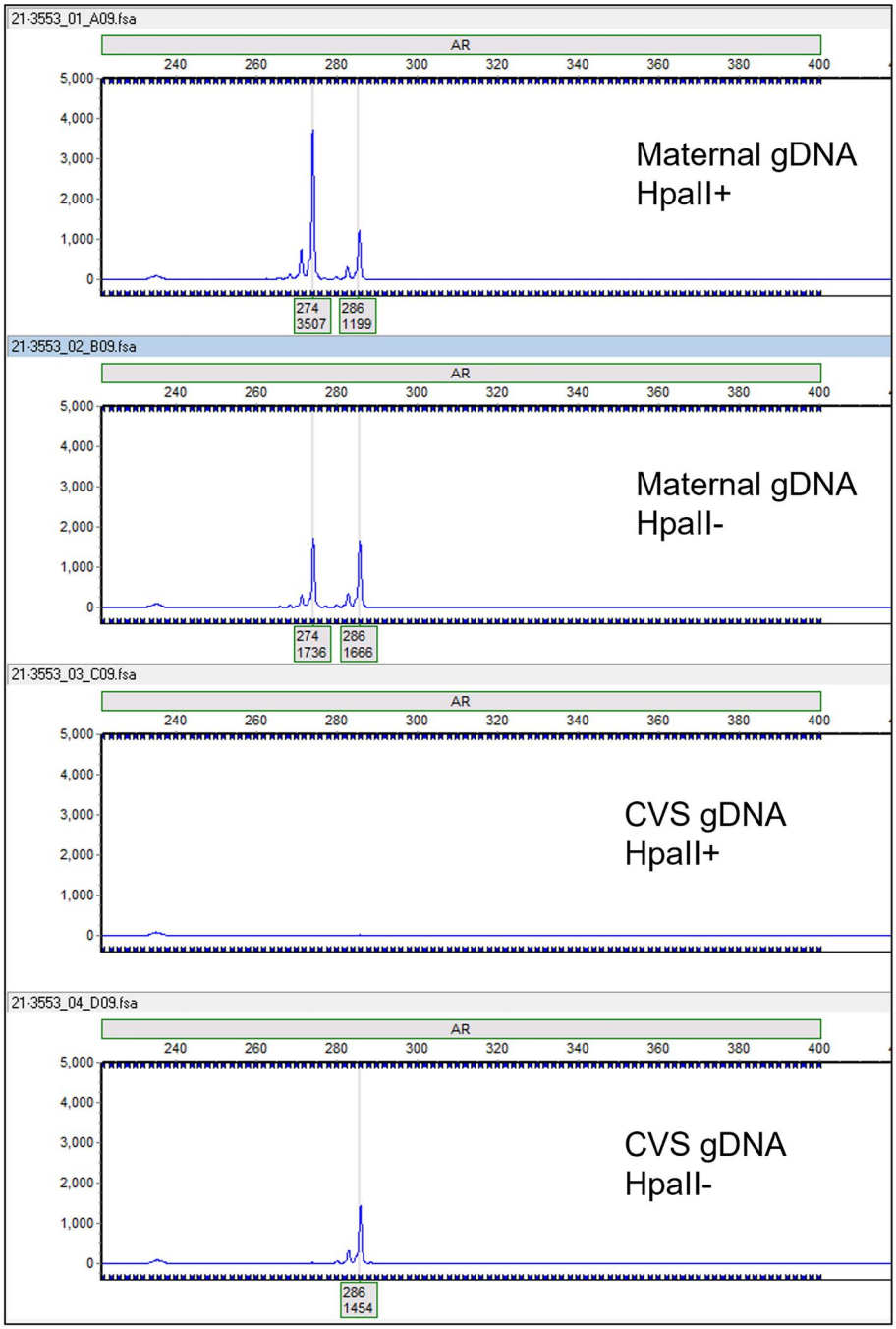

FIGURE 3 X-inactivation results for family 28 (cfDNA-29), viewed in GeneMarker v3.0.1. Peak annotations show the size of each allele in bp and peak height in relative fluorescence units. The male fetus inherited the X chromosome with a 286bp *AR* amplicon, which was preferentially unmethylated in maternal genomic DNA (gDNA), as shown by the skewed peak heights when digested with the methylation-sensitive enzyme HpaII. This X chromosome did not contain the *IDS* c.182_189del pathogenic variant. [Colour figure can be viewed at wileyonlinelibrary.com]

conclusively explain the aetiology of either incorrect prediction. A sequence polymorphism within the probe or primer binding sites could have caused the incorrect results by affecting the amplification rate. However, Sanger sequencing showed no sequence variants within the primer and probe binding sites for either assay. The patients' medical notes were also reviewed: neither patient had a blood transfusion prior to pregnancy or prior to the test, neither patient had an IVF pregnancy, and there was no report of another gestational sac on the dating scan ultrasound reports, indicating that the incorrect predictions were not caused by a vanishing twin.

The false positive result for the *IDS* c.182_189del variant may have been due to the impact of XCI on cfDNA fragment length and therefore amplifiability using ddPCR. One hypothesis is that increased activation and reduced methylation of the X chromosome bearing the normal allele (Xn) would lead to the production of shorter cfDNA fragments, as endonuclease enzymes have greater access to

DNA in less condensed chromatin. Shorter cfDNA fragments would be less likely to be amplifiable by ddPCR, due to the decreased probability of both primer-binding sequences being present within a single fragment. If there were fewer amplifiable fragments for the normal allele, this could make it appear that the pathogenic variant was higher in abundance, leading to the incorrect fetal genotype prediction. However, this hypothesis has not been confirmed. Notably, the 8bp *IDS* c.182_189del deletion (CCCCAAAT) contains two common 4-mer end motifs (CCCA and CCAA) reported to be generated by DNASE1L3 digestion during the production of cfDNA,[40] which may have implications for the fragmentation of this cfDNA amplicon. This hypothesis requires further investigation; however, this is not within the scope of this study since we do not have access to pre-pregnancy samples. In a clinical setting, before testing for X-linked conditions a maternal sample taken prior to the pregnancy would be required in order to check the allelic balance of the assay. If an imbalance in the variant fraction is identified, we would currently advise against non-invasive prenatal diagnosis until this XCI hypothesis can be further investigated.

While this is a large retrospective cohort study of RMD with ddPCR, it has several limitations. Firstly, this approach was not a comprehensive comparison of all available methods. For example, we did not include the chi-squared and Bayes factor approach reported previously by Chang et al.[26] as this requires the generation of multiple biological control samples for each family, which is neither economical for a large sample cohort nor in clinical service. Secondly, whilst using results from gDNA control samples as a reference dataset is an improvement on statistical assumptions, this still has limitations and by accounting for technical variation in the assay the rate of incorrect results was reduced but not eliminated. The results of testing heterozygous sonicated gDNA showed no systematic bias of any assays toward either allele (Supplementary Figure 2), indicating that assay validation was not a contributing factor to the incorrect results. However, cfDNA has many unique features, including a distinct fragment size profile, jaggedness, non-random end-motifs and topology,[41] which may impact ddPCR assays and cannot be replicated using gDNA as a surrogate reference material. Unfortunately, we did not have access to blood samples taken prior to pregnancy for the women in our cohort, which would have been the most appropriate control material. By applying a cfDNA reference dataset instead of gDNA controls, the incorrect results with z-score analysis may have been prevented, which suggests an avenue for future development. In an ideal scenario, each assay would have been optimised on maternal cfDNA from a pre-pregnancy control plasma sample to assess for any allelic imbalances. A further limitation is that we used archived samples, some of which had low fetal fraction, and therefore we could not access repeat samples later in pregnancy when fetal fraction may be higher.

Recent reports have applied NGS to RMD, with a focus on SCD, with modifications to allow for single-molecule counting.[22,23,42] These approaches allow sequence-level inspection of the cfDNA and could theoretically be applied to the detection of any pathogenic variant. However, ddPCR is a cheaper and more flexible technique, which allows the rapid development of assays for both rare and common variants, giving it a significant advantage over NGS techniques for clinical application.

5 | CONCLUSION

In summary, we report a large cohort of cfDNA analysis for maternally inherited variants, including the common *HBB* c.20A>T variant. In our cohort of 124 cfDNA samples, z-score analysis was the most accurate method, generating 98% correct fetal genotype predictions out of the total number of predictions made, but with a high inconclusive rate of 31%. However, based on the evidence from previous NIPD tests, which have been translated into clinical service, this inconclusive rate would be expected to reduce with repeat sampling. For the two incorrect fetal genotype predictions made by the z-score analysis, the cause remains unknown. It may be that with further refinements to the analysis method it will become possible to reduce this rate of incorrect results. High degrees of accuracy are required prior to clinical implementation of cfDNA tests, and clear information on test performance must be given to patients and physicians.

ACKNOWLEDGEMENTS

We wish to acknowledge Dr. Joan Camunas-Soler for providing additional details of the ddPCR SNP panel used for fetal fraction determination and Bonnie Luxford for her contributions to ddPCR for the dominant inheritance cases.

This work is funded by the National Institute for Health Research Great Ormond Street Hospital Biomedical Research Centre (NIHR GOSH BRC). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

This research is funded by the NIHR Biomedical Research Centre at Great Ormond Street Hospital, London, UK.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

CONFLICT OF INTEREST STATEMENT

The authors declare that there is no conflict of interest.

ORCID

Joe Shaw https://orcid.org/0000-0002-5089-751X
Elizabeth Scotchman https://orcid.org/0000-0001-7418-6037
Lyn S. Chitty https://orcid.org/0000-0002-4857-7138
Natalie Chandler https://orcid.org/0000-0003-1396-0740

REFERENCES

1. Lo YM, Corbetta N, Chamberlain PF, et al. Presence of fetal DNA in maternal plasma and serum. *Lancet*. 1997;350(9076):485-487. https://doi.org/10.1016/s0140-6736(97)02174-0

2. Chiu RW, Chan KC, Gao Y, et al. Noninvasive prenatal diagnosis of fetal chromosomal aneuploidy by massively parallel genomic sequencing of DNA in maternal plasma. *Proc Natl Acad Sci U S A*. 2008; 105(51):20458-20463. https://doi.org/10.1073/pnas.0810641105
3. Devaney SA, Palomaki GE, Scott JA, Bianchi DW. Noninvasive fetal sex determination using cell-free fetal DNA: a systematic review and meta-analysis. *JAMA*. 2011;306(6):627-636. https://doi.org/10.1001/jama.2011.1114
4. Lo YM, Hjelm NM, Fidler C, et al. Prenatal diagnosis of fetal RhD status by molecular analysis of maternal plasma. *N Engl J Med*. 1998;339(24):1734-1738. https://doi.org/10.1056/nejm199812103392402
5. Drury S, Mason S, Mckay F, et al. Implementing non-invasive prenatal diagnosis (NIPD) in a national health service laboratory; from dominant to recessive disorders. *Adv Exp Med Biol*. 2016;924:71-75.
6. Lo YM, Chan KC, Sun H, et al. Maternal plasma DNA sequencing reveals the genome-wide genetic and mutational profile of the fetus. *Sci Transl Med*. 2010;2(61):61ra91. https://doi.org/10.1126/scitranslmed.3001720
7. Piel FB, Hay SI, Gupta S, Weatherall DJ, Williams TN. Global burden of sickle cell anaemia in children under five, 2010-2050: modelling based on demographics, excess mortality, and interventions. *PLoS Med*. 2013;10(7):e1001484. https://doi.org/10.1371/journal.pmed.1001484
8. Telfer P, Coen P, Chakravorty S, et al. Clinical outcomes in children with sickle cell disease living in England: a neonatal cohort in East London. *Haematologica*. 2007;92(7):905-912. https://doi.org/10.3324/haematol.10937
9. *NHS Sickle Cell and Thalassaemia Screening Programme Data Report 2017 to 2018*. Public Health England; 2020.
10. Young E, Bowns B, Gerrish A, et al. Clinical service delivery of noninvasive prenatal diagnosis by relative haplotype dosage for single-gene disorders. *J Mol Diagn*. 2020;22(9):1151-1161. https://doi.org/10.1016/j.jmoldx.2020.06.001
11. Hill M, Oteng-Ntim E, Forya F, Petrou M, Morris S, Chitty LS. Preferences for prenatal diagnosis of sickle-cell disorder: a discrete choice experiment comparing potential service users and healthcare providers. *Health Expect*. 2017;20(6):1289-1295. https://doi.org/10.1111/hex.12568
12. Lun FM, Tsui NB, Chan KC, et al. Noninvasive prenatal diagnosis of monogenic diseases by digital size selection and relative mutation dosage on DNA in maternal plasma. *Proc Natl Acad Sci U S A*. 2008; 105(50):19920-19925. https://doi.org/10.1073/pnas.0810373105
13. Barrett AN, Zimmermann BG, Wang D, Holloway A, Chitty LS. Implementing prenatal diagnosis based on cell-free fetal DNA: accurate identification of factors affecting fetal DNA yield. *PLoS One*. 2011;6(10):e25202. https://doi.org/10.1371/journal.pone.0025202
14. Lo YM, Tein MS, Lau TK, et al. Quantitative analysis of fetal DNA in maternal plasma and serum: implications for noninvasive prenatal diagnosis. *Am J Hum Genet*. 1998;62(4):768-775. https://doi.org/10.1086/301800
15. Tsui NB, Kadir RA, Chan KC, et al. Noninvasive prenatal diagnosis of hemophilia by microfluidics digital PCR analysis of maternal plasma DNA. *Blood*. 2011;117(13):3684-3691. https://doi.org/10.1182/blood-2010-10-310789
16. Hudecova I, Jiang P, Davies J, Lo YMD, Kadir RA, Chiu RWK. Noninvasive detection of F8 int22h-related inversions and sequence variants in maternal plasma of hemophilia carriers. *Blood*. 2017; 130(3):340-347. https://doi.org/10.1182/blood-2016-12-755017
17. Caswell RC, Snowsill T, Houghton JL, et al. Noninvasive fetal genotyping by droplet digital PCR to identify maternally inherited monogenic diabetes variants. *Clin Chem*. 2020;66(7):958-965. https://doi.org/10.1093/clinchem/hvaa104
18. Chang MY, Kim AR, Kim MY, et al. Development of novel noninvasive prenatal testing protocol for whole autosomal recessive disease using picodroplet digital PCR. *Sci Rep*. 2016;6(1):37153. https://doi.org/10.1038/srep37153
19. Barrett AN, Mcdonnell TC, Chan KC, Chitty LS. Digital PCR analysis of maternal plasma for noninvasive detection of sickle cell anemia. *Clin Chem*. 2012;58(6):1026-1032. https://doi.org/10.1373/clinchem.2011.178939
20. Sawakwongpra K, Tangmansakulchai K, Ngonsawan W, et al. Droplet-based digital PCR for non-invasive prenatal genetic diagnosis of α and β-thalassemia. *Biomed Rep*. 2021;15(4):82. https://doi.org/10.3892/br.2021.1458
21. Camunas-Soler J, Lee H, Hudgins L, et al. Noninvasive prenatal diagnosis of single-gene disorders by use of droplet digital PCR. *Clin Chem*. 2018;64(2):336-345. https://doi.org/10.1373/clinchem.2017.278101
22. Tsao DS, Silas S, Landry BP, et al. A novel high-throughput molecular counting method with single base-pair resolution enables accurate single-gene NIPT. *Sci Rep*. 2019;9(1):14382. https://doi.org/10.1038/s41598-019-50378-8
23. Van Campen J, Silcock L, Yau S, et al. A novel non-invasive prenatal sickle cell disease test for all at-risk pregnancies. *Br J Haematol*. 2020;190(1):119-124. https://doi.org/10.1111/bjh.16529
24. Perlado S, Bustamante-Aragones A, Donas M, Lorda-Sanchez I, Plaza J, Rodriguez de Alba M. Fetal genotyping in maternal blood by digital PCR: towards NIPD of monogenic disorders independently of parental origin. *PLoS One*. 2016;11(4):e0153258. https://doi.org/10.1371/journal.pone.0153258
25. Gu W, Koh W, Blumenfeld YJ, et al. Noninvasive prenatal diagnosis in a fetus at risk for methylmalonic acidemia. *Genet Med*. 2014; 16:564-567. https://doi.org/10.1038/gim.2013.194
26. Chang MY, Ahn S, Kim MY, et al. One-step noninvasive prenatal testing (NIPT) for autosomal recessive homozygous point mutations using digital PCR. *Sci Rep*. 2018;8(1):2877. https://doi.org/10.1038/s41598-018-21236-w
27. Orhant L, Anselem O, Fradin M, et al. Droplet digital PCR combined with minisequencing, a new approach to analyze fetal DNA from maternal blood: application to the non-invasive prenatal diagnosis of achondroplasia. *Prenat Diagn*. 2016;36(5):397-406. https://doi.org/10.1002/pd.4790
28. Gruber A, Pacault M, El Khattabi LA, et al. Non-invasive prenatal diagnosis of paternally inherited disorders from maternal plasma: detection of NF1 and CFTR mutations using droplet digital PCR. *Clin Chem Lab Med*. 2018;56(5):728-738. https://doi.org/10.1515/cclm-2017-0689
29. Lun FM, Chiu RW, Chan KC, Yeung Leung T, Kin Lau T, Dennis Lo YM. Microfluidics digital PCR reveals a higher than expected fraction of fetal DNA in maternal plasma. *Clin Chem*. 2008;54(10): 1664-1672. https://doi.org/10.1373/clinchem.2008.111385
30. Preka E, Ellershaw D, Chandler N, et al. Cell-free DNA in pediatric solid organ transplantation using a new detection method of separating donor-derived from recipient cell-free DNA. *Clin Chem*. 2020;66(10):1300-1309. https://doi.org/10.1093/clinchem/hvaa173
31. Lo YM, Lun FM, Chan KC, et al. Digital PCR for the molecular detection of fetal chromosomal aneuploidy. *Proc Natl Acad Sci U S A*. 2007; 104(32):13116-13121. https://doi.org/10.1073/pnas.0705765104
32. Royall R. *Statistical Evidence: A Likelihood Paradigm*. Chapman & Hall; 1997:179.
33. Canick JA, Palomaki GE, Kloza EM, Lambert-Messerlian GM, Haddow JE. The impact of maternal plasma DNA fetal fraction on next generation sequencing tests for common fetal aneuploidies. *Prenat Diagn*. 2013;33(7):667-674. https://doi.org/10.1002/pd.4126

34. Deng C, Liu S. Factors affecting the fetal fraction in noninvasive prenatal screening: a review. *Front Pediatr*. 2022;10:812781. https://doi.org/10.3389/fped.2022.812781
35. Lun FM, Chiu RW, Sun K, et al. Noninvasive prenatal methylomic analysis by genomewide bisulfite sequencing of maternal plasma DNA. *Clin Chem*. 2013;59(11):1583-1594. https://doi.org/10.1373/clinchem.2013.212274
36. Chandler NJ, Ahlfors H, Drury S, et al. Noninvasive prenatal diagnosis for cystic fibrosis: implementation, uptake, outcome, and implications. *Clin Chem*. 2020;66(1):207-216. https://doi.org/10.1373/clinchem.2019.305011
37. Wald A. *Sequential Analysis*. John Wiley & Sons; 1947:224.
38. Xiong L, Barrett AN, Hua R, et al. Non-invasive prenatal testing for fetal inheritance of maternal beta-thalassaemia mutations using targeted sequencing and relative mutation dosage: a feasibility study. *Bjog*. 2018;125(4):461-468. https://doi.org/10.1111/1471-0528.15045
39. El Karoui N, Zhou W, Whittemore AS. Getting more from digital SNP data. *Stat Med*. 2006;25(18):3124-3133. https://doi.org/10.1002/sim.2379
40. Serpas L, Chan RWY, Jiang P, et al. Dnase1l3 deletion causes aberrations in length and end-motif frequencies in plasma DNA. *Proc Natl Acad Sci U S A*. 2019;116(2):641-649. https://doi.org/10.1073/pnas.1815031116
41. Lo YMD, Han DSC, Jiang P, Chiu RWK. Epigenetics, fragmentomics, and topology of cell-free DNA in liquid biopsies. *Science*. 2021;372(6538):372. https://doi.org/10.1126/science.aaw3616
42. Cutts A, Vavoulis DV, Petrou M, et al. A method for non-invasive prenatal diagnosis of monogenic autosomal recessive disorders. *Blood*. 2019;134(14):1190-1193. https://doi.org/10.1182/blood.2019002099

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Shaw J, Scotchman E, Paternoster B, et al. Non-invasive fetal genotyping for maternal alleles with droplet digital PCR: A comparative study of analytical approaches. *Prenat Diagn*. 2023;43(4):477-488. https://doi.org/10.1002/pd.6333