



# Variability in Fetal Fraction Estimation: Comparing Fetal Fractions Reported by Noninvasive Prenatal Testing Providers Globally

Ellis C. Becking <sup>a</sup>, Jasper Linthorst,<sup>b,c</sup> Simon Patton,<sup>d</sup> Weronika Gutowska-Ding,<sup>d</sup> Rebecca Goodall,<sup>d</sup> Farrah Khawaja,<sup>e</sup> Fiona Morgan,<sup>e</sup> Zandra Deans,<sup>e</sup> Lyn S. Chitty,<sup>f</sup> Mireille N. Bekker,<sup>a</sup> Peter G. Scheffer,<sup>a</sup> and Erik A. Sistermans <sup>b,c,\*</sup>

**BACKGROUND:** Fetal fraction (FF) measurement is considered important for reliable noninvasive prenatal testing (NIPT). Using minimal FF threshold as a quality parameter is under debate. We evaluated the variability in reported FFs of individual samples between providers and laboratories and within a single laboratory.

**METHODS:** Genomic quality assessment and European Molecular Genetics Quality Network provide joint proficiency testing for NIPT. We compared reported FFs across all laboratories and stratified according to test methodologies. A single sample was sequenced repeatedly and FF estimated by 2 bioinformatics methods: Veriseq2 and SeqFF. Finally, we compared FFs by Veriseq and SeqFF in 87 351 NIPT samples.

**RESULTS:** For each proficiency test sample we observed a large variability in reported FF, SDs and CVs ranging from 1.7% to 3.6% and 17.0% to 35.8%, respectively. FF measurements reported by single nucleotide polymorphism-based methods had smaller SDs (0.5% to 2.4%) compared to whole genome sequencing-based methods (1.8% to 2.9%). In the internal quality assessment, SDs were similar between SeqFF (SD = 1.0%) and Veriseq v2 (SD = 0.9%), but mean FF by Veriseq v2 was higher compared to SeqFF (9.0% vs 6.4%,  $P < 0.001$ ). In patient samples, reported FFs were on average 1.12%-points higher in Veriseq than in SeqFF ( $P < 0.001$ ).

**CONCLUSIONS:** Current methods do not allow for a reliable and consistent FF estimation. Our data show estimated FF should be regarded as a laboratory-specific

range, rather than a precise number. Applying strict universal minimum thresholds might result in unnecessary test failures and should be used with caution.

## Introduction

The presence of cell-free placental DNA, often referred to as cell-free fetal DNA, in maternal blood plasma enables early noninvasive prenatal testing (NIPT) for fetal aneuploidies (1). The performance of these tests depends on a sufficiently large amount of cell-free fetal DNA, referred to as fetal fraction (FF), which is estimated in most NIPT methods and serves as a quality control parameter. FF can vary based on gestational age, maternal weight, and other biological factors of which most are currently unknown (2–5).

There are basically 2 ways in which FF estimates can be obtained independent of the sex of the fetus. Single nucleotide polymorphism (SNP)-based methods calculate the abundance of different alleles at specifically targeted common SNPs (6). Whole genome sequencing (WGS)-based methods use various characteristics of the sequencing data that are known to differ between maternal and fetal cell-free DNA, such as genome-wide distribution and fragment length (2, 7–9). In cases where the fetus is male, the amount of Y-chromosomal cell-free fetal DNA fragments can be used (10).

Although in practice the different approaches to NIPT seem to provide sufficient precision for the

<sup>a</sup>Department of Obstetrics, Division of Women and Baby, Wilhelmina Children's Hospital, University Medical Center Utrecht, Utrecht, The Netherlands; <sup>b</sup>Department of Human Genetics, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands; <sup>c</sup>Amsterdam Reproduction and Development, Amsterdam, The Netherlands; <sup>d</sup>Manchester Science Park, EMQN CIC, Manchester, UK; <sup>e</sup>Genomics Quality Assessment, NHS Lothian, Royal Infirmary of Edinburgh, Edinburgh, UK; <sup>f</sup>North Thames Genomic Laboratory Hub, Great Ormond Street NHS Foundation Trust and Genetics and

Genomics, UCL Great Ormond Street Institute of Child Health, London, UK.

\*Address correspondence to this author at: Department of Human Genetics, Amsterdam UMC, van der Boechorstraat 7, 1081 BT Amsterdam, The Netherlands. E-mail e.sistermans@amsterdamumc.nl. Disclaimer: The views expressed are those of the authors and not necessarily those of the National Health Service, the National Institute of Health Research, or the Department of Health and Social Care. Received August 11, 2022; accepted November 7, 2022. <https://doi.org/10.1093/clinchem/hvac207>

detection of (specific) aneuploidies, it is unclear how accurate reported fetal fractions really are. Previous studies provide evidence of a large variability in reported FF values, based on the method and platform used (11, 12).

NIPT providers often reject samples based on a predetermined low FF cut-off value, with reported “no-call” rates due to low FF in up to 6.1% (median 1.3%) of tests performed (13). Furthermore, FF should also be related to the amount of data generated per individual analysis, as more data will allow for a reliable test at lower FF. This was clearly demonstrated by Persson et al. who recently constructed a model to simulate both estimated FF and true FF and determined the proportion of samples with estimated FF and true FF below various cut-offs (14). This simulation resulted in 64.8% to 99.9% of samples that would be erroneously classified as no-calls, and, although a theoretical approach, it shows that retesting and/or the recommendation of invasive testing based on inadequate FF estimation may lead to unnecessary concern in patients and increased costs. It is therefore important to obtain more insight in the reliability of FF estimation by commonly used NIPT tests between methods, between laboratories, and within a single laboratory.

Here, we examine the variability of FF measurements across different NIPT providers and NIPT approaches. We compared 477 reported FF estimates that were collected in the context of fetal aneuploidy screening by 2 external quality assessment (EQA) providers in 2018, 2019, and 2020: European Molecular Genetics Quality Network (EMQN) and Genomic Quality Assessment (GenQA). In addition, from an internal quality assessment of a single laboratory (Amsterdam), we examine the WGS-based FF estimates in a repeatedly analyzed manufactured trisomy 21 (T21)-positive plasma sample. Finally, we compare 2 methods that were tested in parallel on 87 351 patient samples.

## Materials and Methods

### EMQN AND GENQA NIPT PROFICIENCY TESTING

Each year, for a period of 3 years, 2 cases were provided for aneuploidy screening in a single round. We denote these cases 2018-1, 2018-2, 2019-1, 2019-2, 2020-1, and 2020-2. With the exception of 2018-1, manufactured plasma with a known FF (given in Table 1) as provided by the manufacturer was used for these EQAs. In 2018-1, a range of unique patient samples was used; all had different FFs, and each laboratory received a different sample. Known FFs of provided samples were not revealed to the laboratories until after each round. Laboratories were asked to report according to their normal format; reporting FF was not a prerequisite. The

reported FF was analyzed across all external NIPT providers that included this information in their reports. We further subdivided the reported FF by whether or not an SNP-based approach was used. Based on the information provided to us from the laboratories, all non-SNP-based methods were based on WGS. Within both SNP-based and non-SNP-based methods, a variety of methodologies was used to estimate the FF.

### MANUFACTURED SAMPLES USED FOR PROFICIENCY TESTING

Manufactured samples were provided by Seracare and used according to the recommendations of the supplier. The sample manufacturing process was consistent over the years 2018 to 2020. All samples were custom-made and do not have a catalog and lot number. In agreement with EMQN/GenQA guidelines, all samples were tested and validated by independent laboratories using a variety of NIPT methods before being sent for proficiency testing. In 2018-2 and 2019-1, a sample with a high risk of T21 was used. The sample used in 2020-1 was of a high risk for trisomy 13 (T13). In 2019-2 and 2020-2, samples with a low risk of T13, trisomy 18 (T18), and T21 with evidence of XXY were used. All samples were fetally male. All laboratories received 1 mL of each sample (though, in 2018, a 4 mL sample was used) and repeat samples could be obtained upon request. The patient samples of the 2018-1 scheme were provided by the RAPID sample bank based at Great Ormond Street Hospital in London. Maternal blood was collected into EDTA tubes if DNA extraction could occur within 24 h of blood draw, or cell stabilizing tubes if not, and cell-free DNA was extracted as described previously (15).

### PATIENT SAMPLES USED FOR PROFICIENCY TESTING

On receipt, samples were booked in with a unique Biobank number, ensuring that the study ID and hospital unit code on each sample and referral form matched. All transfers were performed in a class 2 safety cabinet and had to be checked by a second operator. Samples received in glass Streck tubes could not be centrifuged and were transferred to 15 mL falcon tubes, each labeled with a biobank number and study ID/hospital unit number. Samples were spun at 1500 rcf (g) for 10 min at room temperature using a slow brake and acceleration speed. The maternal plasma was then aspirated into labeled 1.5 mL Eppendorf (Lobind) tubes using a sterile 2 mL plastic pastette and spun at 16 000 rcf (g) for 10 min using brake and acceleration settings 9 (the shortest run up and run down time). The blood pellets were pooled into EDTA/monovette tubes labeled with biobank stickers. The plasma was then transferred into the correctly labeled 2 mL Eppendorf (Lobind) tubes using a sterile 2 mL plastic

**Table 1. Summary of the analysis of reported fetal fractions from NIPT proficiency testing 2018–2020.**

	2018		2019		2020			
	2018-1 <sup>a</sup>	2018-2	2019-1	2019-2	2020-1	2020-2		
Samples sent out for analysis (n)	73	69	100	103	121	121		
Results received for FF [n (%)]	61 (83.6)	58 (84.1)	88 (88)	92 (89.3)	89 (73.6)	89 (73.6)		
Methodology used (n)	NGS	40	NGS	38	NGS	75	NGS	74
	Array	7	Array	7	Array	3	Array	7
	PCR	0	PCR	0	PCR	1	PCR	1
	Unknown	14	Unknown	13	Unknown	10	Unknown	9
FF reported by provider (%)	Variable <sup>b</sup>	4	12	16	12	11		
	Reported FF (%)		Reported FF (%)		Reported FF (%)			
Median	9.2	7.3	7.8	16.9	13.5	9		
Mean	10.3	7.8	8.5	16.3	13.6	10.2		
Minimum	4	5	4.1	10	5.6	6		
Maximum	18	12.7	14	24.2	28	26.8		
Interquartile range	4	1.6	2.3	3.6	2.3	2		
SD	3.6	1.7	2.7	2.6	2.3	3.6		
CV (%)	34.7	21.6	31.2	30.9	17.0	35.8		
Median deviation from manufacturer reported FF	NA	3.3	4.2	2	1.9	2.6		

<sup>a</sup>All samples were manufactured with the exception of 2018-1.  
<sup>b</sup>In this scheme different samples with different FFs were used.  
NA, not applicable; NGS, next generation sequencing.

pastette. Plasma aliquots and blood pellets were stored in a designated biobank  $-80^{\circ}\text{C}$  freezer (16).

#### INTERNAL QUALITY ASSESSMENT

The Amsterdam NIPT laboratory used a T21-positive sample obtained from the same provider (Seracare) as an internal quality control in every 96-sample run. During this period, this laboratory used the Veriseq v2 system provided by Illumina. These data were used to determine the consistency of WGS-based FF estimates using the same protocol, staff, and equipment. A total of 177 of these positive controls with an expected FF of 5.7% as stated by the provider were sequenced and analyzed. FFs were estimated by both Veriseq v2 and SeqFF (10). Veriseq v2 documentation states that it combines information from the distribution of both the lengths and genomic coordinates of the library fragments to make FF predictions (17). SeqFF was developed to make FF predictions independent of fragment size information and applies a read-count based model that uses read counts in 50 kb bins across the genome to estimate the FF (8).

To further examine reported FF by Veriseq (v1 and v2) and SeqFF, we compared FF predictions in a larger

series of 87 351 patient samples originating from the Amsterdam NIPT laboratory.

#### STATISTICAL ANALYSIS

Descriptive statistics were provided for all cases. Outliers were defined as 1.5 times the interquartile range from the 1st and 3rd quartile. The CV was reported for case 2018-1 to 2020-2 and defined as the SD relative to the mean FF and is expressed as a percentage. In addition, we report the deviation from the manufacturer reported FF (the absolute difference between the reported FF by the NIPT laboratory and the FF in the manufactured plasma stated by the manufacturer).

When applicable, the difference in mean reported FF was compared using the 2-sample *t*-test. To compare variability (variance) between methods, a Levene test was used (18). In the EMQN and GenQA NIPT proficiency testing, FF was first transformed using mean-normalization per case, to ensure comparability across all cases. Case 2018-1 was excluded from this analysis, because different patient samples with varying fractions were used. Correlations between measurements were calculated by Pearson correlation coefficient (Pearson *r*). All statistical analyses were performed in RStudio (version 1.3.1093).

Results

REPORTING FETAL FRACTION

Participating laboratories in the NIPT proficiency testing reported the FF in integers or one to 2 decimal places. In the internal quality assessment, SeqFF reported the FF in 6 decimal places, while the FF estimated by Veriseq v2 was reported in integers.

EMQN AND GENQA NIPT PROFICIENCY TESTING

A summary of the analysis of the reported FF is provided in Table 1. From 2018 to 2020, a total of 587 samples were sent out to the NIPT providing laboratories. The majority of participating laboratories from 2018 to 2020 originated from Europe (77.3%), followed by Asia (11.2%), Australia (6.0%), South America (2.7%), North America (1.7%), and Africa (1.1%). Laboratories were allowed to request repeat samples when the initial analysis failed. The percentage of laboratories submitting an estimated FF ranged from 73.6% to 89.3%. Due to the COVID-19 pandemic, response rates were lower in 2020 compared to other years.

Of all samples analyzed with a reported methodology by the laboratory (n = 415), 90.6% of samples were analyzed using some form of next generation sequencing technology, 8.9% of samples were analyzed using an array-based approach, and in 0.5%, digital PCR was used. For 62 samples, the methodology was not reported.

The variability in the reported fetal fractions per case are presented in Fig. 1. In all cases, a large variability in reported FF was observed. In some cases, the single manufacturer sample performed comparable to the mix of samples sent in case 2018-1. As an example, the interquartile range of 2019-2 was 3.6% vs 4.0% for 2018-1, and the SD of 2020-2 was the same as that of 2018-1 (both 3.6%). The CV was 34.7% in case 2018-1 and had a wide range between cases 2018-2 to 2020-2 (17.0% to 35.8%). With the exception of 2019-2, the mean FF reported by the laboratories differed by a significant and large amount from the FF provided by the manufacturer. Equal FF in the manufactured plasma (12% in both 2019-1 and 2020-1) resulted in an underestimation in 2019-1 and an overestimation in 2020-2.

We subdivided the reported FF based on whether or not an SNP-based methodology was used (Table 2 and Fig. 2). Information on whether or not SNP information was used in the FF estimation was unavailable for 26 to 33 laboratories per case. The SD of SNP-based methods was significantly lower compared to WGS-based methods ( $P=0.006$ ; Levene test), suggesting a more consistent FF measurement across SNP-based methods. However, SNP-based methods did not consistently result in a smaller deviation from the manufacturer reported FF; in 2 cases (2019-1 and 2020-2), the median deviation was even higher

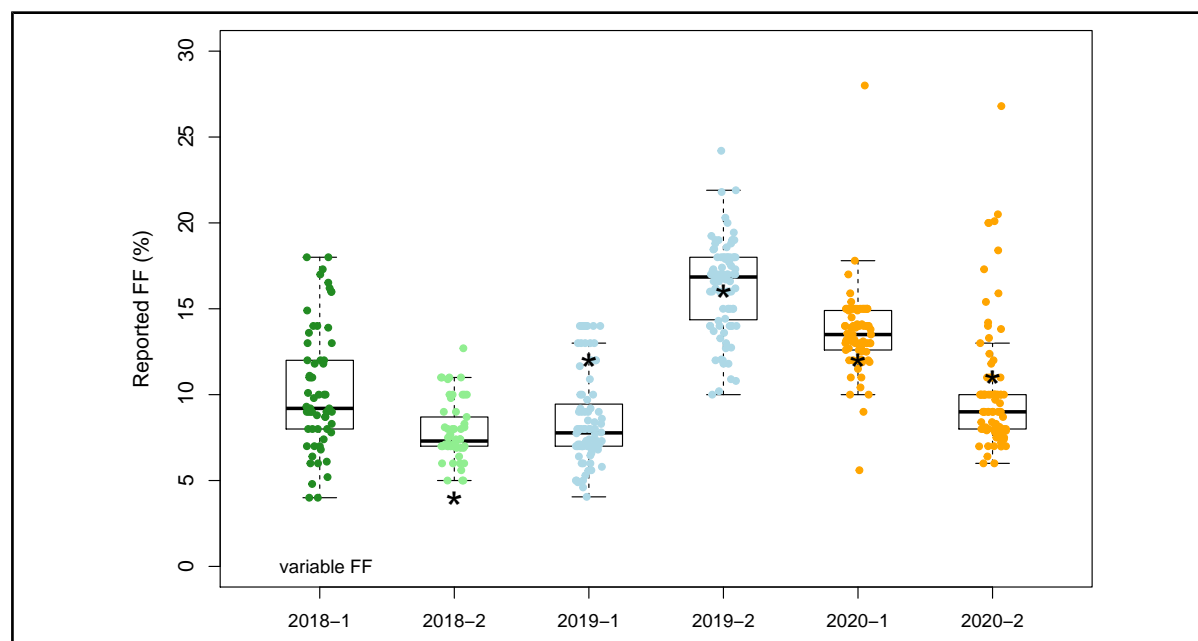
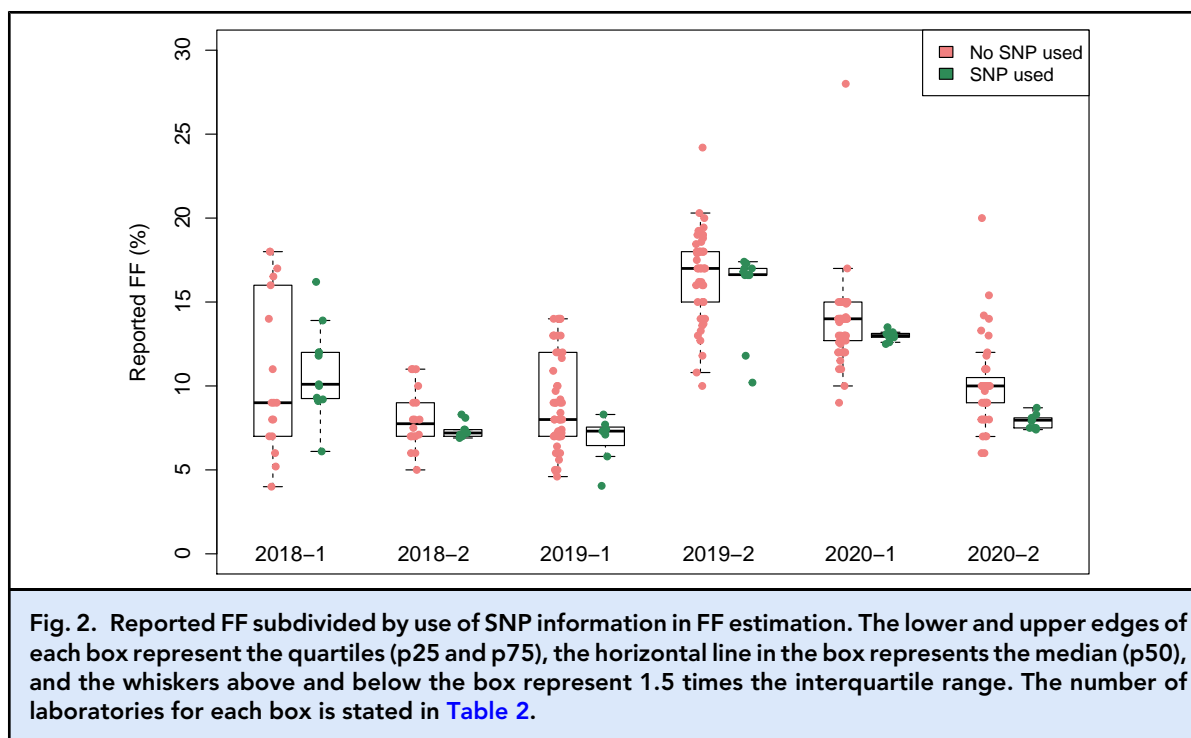


Fig. 1. Distribution of the reported FF (%) in NIPT proficiency testing from cases 2018–2020. The asterisk represents the FF as stated by the manufacturer (not applicable for 2018-1). The lower and upper edges of each box represent the quartiles (p25 and p75), the horizontal line in the box represents the median (p50), and the whiskers above and below the box represent 1.5 times the interquartile range.

**Table 2. Summary of the analysis of reported fetal fractions by use of SNP information in FF estimation.**

	2018				2019				2020			
	2018-1		2018-2		2019-1		2019-2		2020-1		2020-2	
	SNP-based method used				SNP-based method used				SNP-based method used			
	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Laboratories (n)	11	17	9	18	7	54	11	53	9	54	10	48
SD	2.7	4.8	0.5	1.8	1.4	2.9	2.4	2.6	0.3	2.4	0.4	2.4
Median deviation from manufacturer reported FF	NA <sup>a</sup>	NA <sup>a</sup>	3.2	3.8	4.7	4	1	2	1	2	3	1

<sup>a</sup>Not applicable.



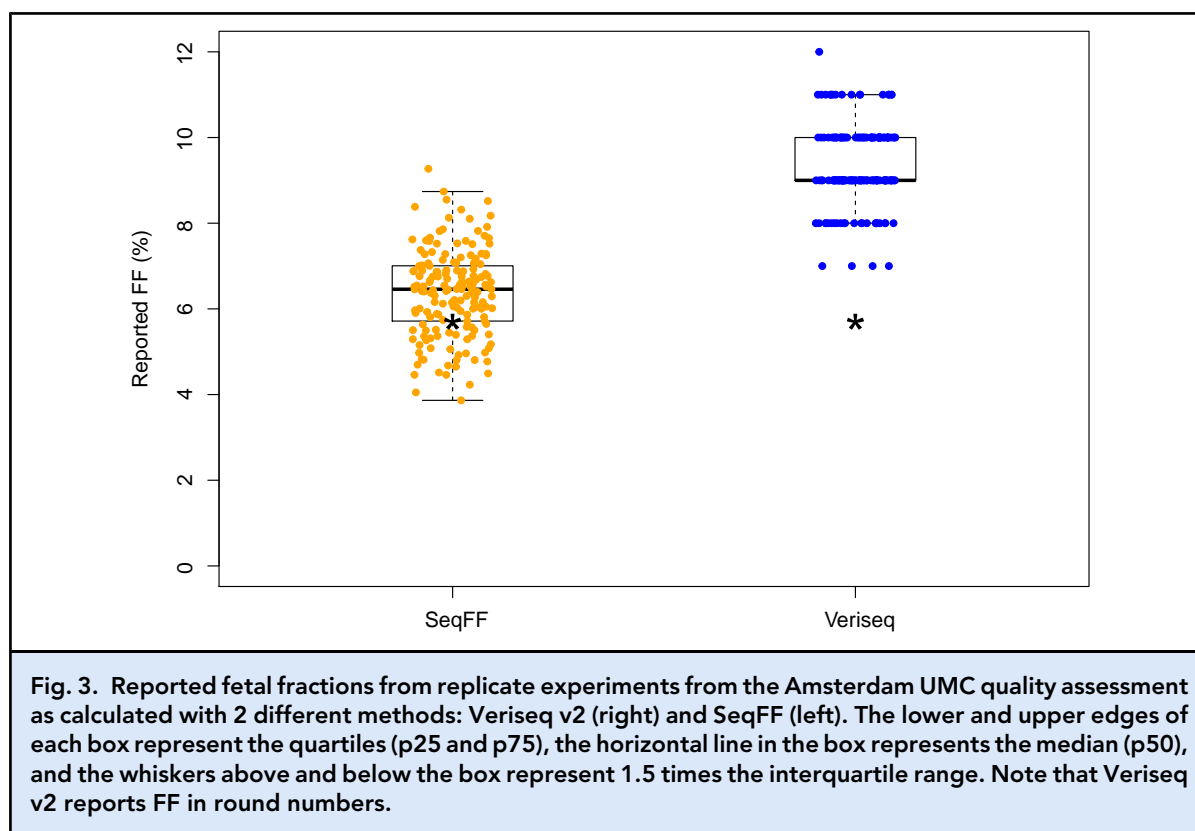
compared to WGS-based methods (Fig. 2). Within laboratories applying SNP-based methods, we did not find significant differences in FF estimations between providers applying next generation sequencing-based vs array-based technologies in any of the cases ( $P$  value  $> 0.05$ ).

#### INTERNAL QUALITY ASSESSMENT (AMSTERDAM UNIVERSITY MEDICAL CENTERS)

The results of the internal quality assessment of the Amsterdam NIPT laboratory are displayed in Fig. 3. Identical T21-positive Seracare samples with a FF of

5.7% (according to the manufacturer) were analyzed 177 times in independent runs. The FF was reported in 171 (96.6%) cases using SeqFF and in 169 cases using Veriseq v2 (95.5%).

For SeqFF, FF ranged from 3.9% to 9.3% with a mean of 6.4% and a median of 6.5%. Mean and median reported FF in the Veriseq v2 analysis were 9.3% and 9.0%, respectively, and FF ranged from 7.0% to 12.0%. On average, reported FF by Veriseq v2 was higher compared to FF predictions by SeqFF ( $t$ -test  $P < 0.001$ ). The variance in reported FF was comparable between both methods, with an SD of 1.0 and 0.9 for SeqFF and Veriseq v2, respectively (Levene test



$P=0.1026$ ). FF measures were moderately correlated (Pearson  $r=0.44$ ). In contrast, the median deviation from manufacturer reported FF for Veriseq v2 was much larger compared to SeqFF (3.3% vs 0.9%). On average, Veriseq v2 estimated the FF of the manufactured plasma to be much higher than the value given by the supplier.

In addition, we examined FF predictions by Veriseq and SeqFF in a series of 87 351 patient samples. Both FF measures were highly correlated (Pearson  $r=0.84$ ), but both the variance and the mean of the Veriseq reported FF were significantly larger (Levene test  $P<0.001$  and  $t$ -test  $P<0.001$ ). Veriseq on average estimated a 1.12%-points higher FF compared to SeqFF (Fig. 4). To generate the bar chart shown in this figure, FF estimates of SeqFF were rounded to the nearest integer before plotting. Samples were excluded from the bar chart if the rounded SeqFF estimate was smaller than 1 or when Veriseq reported the FF as “<1%” or as invalidated.

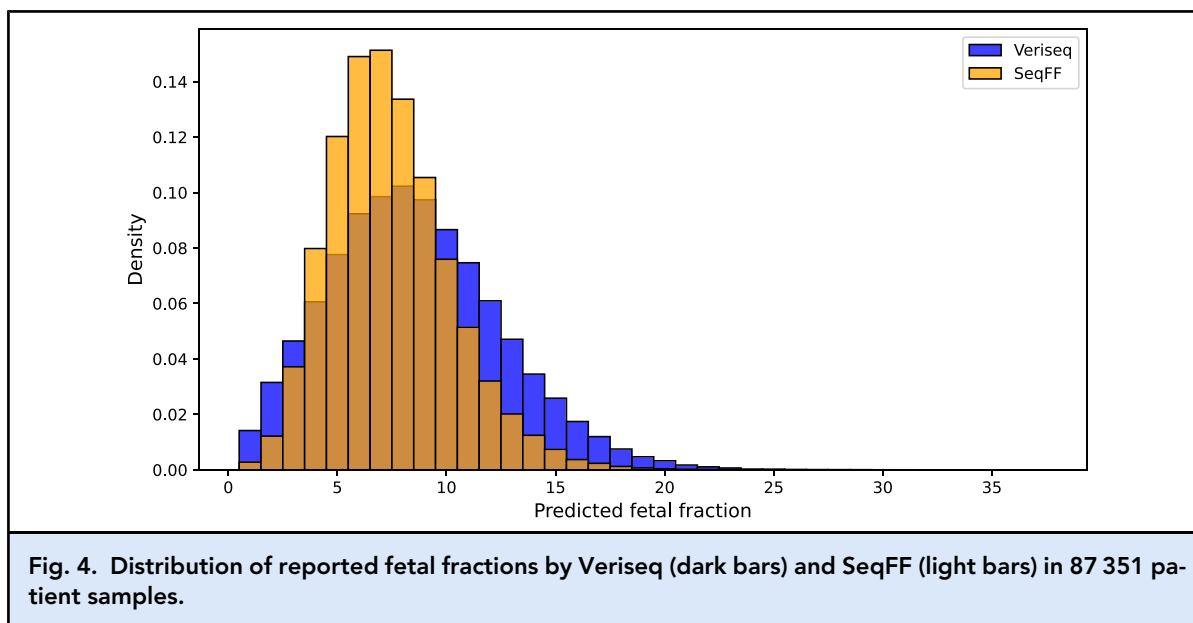
## Discussion

We studied the variability in reported FF across the participants of the NIPT EQAs. Variability in reported FF

was high, especially in methods that were based on WGS. SNP-based methods were found to be more consistent compared to WGS-based methods, but some outliers were still observed. In almost all cases, the mean FF reported differed greatly from the FF stated by the manufacturer.

Compared to the results from NIPT proficiency testing, variability in reported FF was lower in an internal quality assessment with a WGS-based methodology using a single laboratory protocol. The estimated FF was greatly influenced by the type of software used, with Veriseq v2 predicting a mean 2.9% higher FF compared to SeqFF. This suggests that different software implementations for FF prediction account for a sizeable proportion of the observed variance in WGS-based FF predictions.

Increased consistency in FF measurements in SNP-based methods has also been described in a recent study by Persson et al. (11). They identified 6 publications in which the variability of reported FF was obtained and compared between methods used. All investigated methodologies displayed a considerable variability, with SDs ranging from 1.3% to 3.4%. Methods based on SNPs or genomic coordinates and fragment size reported a lower spread (median SD of 1.6) compared to other methods based on nucleosome patterns or fragment size alone. This is confirmed by our results, and the higher



**Fig. 4.** Distribution of reported fetal fractions by Veriseq (dark bars) and SeqFF (light bars) in 87 351 patient samples.

consistency of SNP-based methods is likely to be explained by the fact that these methods directly measure fetal vs maternal SNPs, whereas sequencing-based methods establish FF indirectly based on fragment length and/or differences in genomic coverage. However, there are no indications that SNP-based methods perform better than sequencing-based methods when it comes to trisomy detection, as studies using these methods produce comparable sensitivity, specificity, and positive predictive values.

Despite this imprecision in FF measurements, many laboratories still discard samples based on a too-low FF, usually when the FF is below an arbitrary cut-off of 4% (13, 19). As can be seen from the large spread in our data, overinterpretation of FF and application of these strict cut-offs may result in unnecessary test failures and increases the number of possibly preventable blood redraws or invasive diagnostic procedures.

This raises questions on how to handle high FF variability in clinical practice. This could involve the extension of FF prediction methods to report confidence and/or prediction intervals. Laboratories could also increase assay precision when the FF is reported to be low, for instance by increasing the sequencing depth (20). Some suppliers have already abandoned the use of a strict thresholds by including other relevant quality measures or by specifically increasing FF or targets of interest before analysis.

A strength of this study is that data of a well-organized international multicenter EQA with high reporting were used. We also had access to a large series of patient samples. Our study is limited by the fact that, for a considerable number of NIPT providers, exact

information on methodology was missing, and for some methods it was unclear whether or not SNP information was used. A minor limitation is that all manufactured samples were fetally male, which was done on purpose as it enables participating laboratories to use the Y chromosome-based method to estimate FF. If fetally female samples had also been included, the results could have been even more variable, although most large NIPT providers now use fetal-sex independent estimation methods. The absence of an established gold standard for FF estimation (except possibly for Y chromosome-based methods) makes it difficult to determine a ground truth FF for the samples used during EQA. In almost all experiments, the reported FF is different from the manufacturer, and thus the latter value cannot be used. Laboratories using these samples as positive controls should be aware of this. Another limitation is the use of manufactured samples. The high between-laboratory variability of the estimated FF in manufactured samples may, apart from the applied methodology, to some extent be explained by the use of artificial samples rather than patient samples. To improve reference materials, future research should address how differences between FF estimates in manufactured samples arise. The variance in FF predictions, specifically those from, but not limited to, WGS methods, should be addressed by extending existing prediction methods to report confidence and/or prediction intervals, such that their users are aware of their limitations. Finally, in future years we aim to include samples with low FF (3%–4%), to better study how thresholds are used in practice.

In conclusion, comparison of the results of EQA NIPT samples by many different laboratories using

different methodologies demonstrates a spread in reported FF. Additionally, within a single laboratory this spread is observed at the level of individual tests, and different methods show large differences in FF estimates. Therefore, discarding samples based on low FF alone seems overly conservative and leads to unnecessary blood resampling or even invasive testing, high NIPT costs, and burden for pregnant women. Introduction of new algorithmic developments that jointly model the uncertainties in FF estimates and other parameters into the result of the screening assay should be encouraged to reach the ultimate goal of highly reliable NIPT assays with low failure rates.

**Nonstandard Abbreviations:** NIPT, noninvasive prenatal testing; FF, fetal fraction; SNP, single nucleotide polymorphism; WGS, whole genome sequencing; EQA, external quality assessment; EMQN, European Molecular Genetics Quality Network; GenQA, genomic quality assessment; T21, trisomy 21; T13, trisomy 13; T18, trisomy 18.

**Author Contributions:** All authors confirmed they have contributed to the intellectual content of this paper and have met the following 4 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; (c) final approval of the published article; and (d) agreement to be accountable for all aspects of the article thus

ensuring that questions related to the accuracy or integrity of any part of the article are appropriately investigated and resolved.

**Authors' Disclosures or Potential Conflicts of Interest:** Upon manuscript submission, all authors completed the author disclosure form. Disclosures and/or potential conflicts of interest:

**Employment or Leadership:** E.A. Sistermans is a board member of the society for clinical genetics laboratory diagnostics ("VKGL") in the Netherlands. L.S. Chitty holds a position at the NIHR Great Ormond Street NHS Genomic Medicine Centre.

**Consultant or Advisory Role:** E.A. Sistermans is a member of the advisory board of GenQA.

**Stock Ownership:** None declared.

**Honoraria:** S. Patton, Astrazeneca.

**Research Funding:** Samples for 2018 were obtained from the RAPID sample bank established as part of the National Institute of Health Research (NIHR) Programme Grant for Applied Research (RP-PG-0707-10107). L.S. Chitty is partially funded by the NIHR Biomedical Research Centre at Great Ormond Street Hospital. S. Patton has received financial support for the EQA scheme provision by Amgen, Astrazeneca, and Novartis.

**Expert Testimony:** None declared.

**Patents:** None declared.

**Other Remuneration:** L.S. Chitty, support for attending meetings and/or travel from ISPD.

**Role of Sponsor:** The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, preparation of manuscript, or final approval of manuscript.

## References

- Bianchi DW, Chiu RWK. Sequencing of circulating cell-free DNA during pregnancy. *N Engl J Med* 2018;379:464–73.
- Hestand MS, Bessem M, van Rijn P, de Menezes RX, Sie D, Bakker I, et al. Fetal fraction evaluation in non-invasive prenatal screening (NIPS). *Eur J Hum Genet* 2019; 27:198–202.
- Ashoor G, Poon L, Syngelaki A, Mosimann B, Nicolaides KH. Fetal fraction in maternal plasma cell-free DNA at 11–13 weeks' gestation: effect of maternal and fetal factors. *Fetal Diagn Therapy* 2012;31:237–43.
- Hui L, Bianchi DW. Fetal fraction and non-invasive prenatal testing: what clinicians need to know. *Prenat Diagn* 2020;40:155–63.
- Juul LA, Hartwig TS, Ambye L, Sørensen S, Jørgensen FS. Noninvasive prenatal testing and maternal obesity: a review. *Acta Obstet Gynecol Scand* 2020;99:744–50.
- Dar P, Curnow KJ, Gross SJ, Hall MP, Stosic M, Demko Z, et al. Clinical experience and follow-up with large scale single-nucleotide polymorphism-based non-invasive prenatal aneuploidy testing. *Am J Obstet Gynecol* 2014;211:527.e1–e17.
- Sparks AB, Struble CA, Wang ET, Song K, Oliphant A. Noninvasive prenatal detection and selective analysis of cell-free DNA obtained from maternal blood: evaluation for trisomy 21 and trisomy 18. *Am J Obstet Gynecol* 2012;206:319.e1–e9.
- Kim SK, Hannum G, Geis J, Tynan J, Hogg G, Zhao C, et al. Determination of fetal DNA fraction from the plasma of pregnant women using sequence read counts. *Prenat Diagn* 2015;35:810–5.
- Straver R, Oudejans CBM, Sistermans EA, Reinders MJT. Calculating the fetal fraction for noninvasive prenatal testing based on genome-wide nucleosome profiles. *Prenat Diagn* 2016;36:614–21.
- van Beek DM, Straver R, Weiss MM, Boon EMJ, Huijsdens-van Amsterdam K, Oudejans CBM, et al. Comparing methods for fetal fraction determination and quality control of NIPT samples. *Prenat Diagn* 2017;37:769–73.
- Persson F, Prensly L. Variability of "reported fetal fraction" in noninvasive prenatal screening (NIPS). *Clin Chem* 2021;67:863–6.
- Palomaki GE, Wyatt P, Best RG, Lepage N, Ashwood ER, Souers RJ, et al. Assessment of laboratories offering cell-free (cf) DNA screening for Down syndrome: results of the 2018 College of American Pathology external educational exercises. *Genet Med* 2020;22:777–84.
- Gil MM, Accurti V, Santacruz B, Plana MN, Nicolaides KH. Analysis of cell-free DNA in maternal blood in screening for aneuploidies: updated meta-analysis. *Ultrasound Obstet Gynecol* 2017;50:302–14.
- Persson F, Cuckle HS. Consequences of imprecision in fetal fraction estimation on performance of cell-free DNA screening for Down syndrome. *Prenat Diagn* 2022; 42:512–7.
- Barrett AN, McDonnell TCR, Chan KCA, Chitty LS. Digital PCR analysis of maternal plasma for noninvasive detection of sickle cell anemia. *Clin Chem* 2012;58:1026–32.
- 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:e25202.
- Illumina. VeriSeq NIPT Solution v2 Package Insert (document 1000000078751). [https://support.illumina.com/content/dam/illumina-support/documents/documentation/chemistry\\_documentation/veriseq-nipt-v2/1000000078751\\_08\\_veriseq-nipt-solution-v2-package-insert.pdf](https://support.illumina.com/content/dam/illumina-support/documents/documentation/chemistry_documentation/veriseq-nipt-v2/1000000078751_08_veriseq-nipt-solution-v2-package-insert.pdf) (Accessed December 2022).
- Derrick B, Candidate P, Ruck A, Toher D, White P. Tests for equality of variances between two samples which contain both paired observations and independent observations. *J Appl Quant Methods* 2018; 13:36–47.
- Gregg AR, Skotko BG, Benkendorf JL, Monaghan KG, Bajaj K, Best RG, et al. Noninvasive prenatal screening for fetal aneuploidy, 2016 update: a position statement of the American College of Medical Genetics and Genomics. *Genet Med* 2016;18:1056–65.
- Miceikaitė I, Brasch-Andersen C, Fagerberg C, Larsen MJ. Total number of reads affects the accuracy of fetal fraction estimates in NIPT. *Mol Genet Genomic Med* 2021;9:e1653.