Received: 8 September 2020 | Revised: 2 February 2021 | Accepted: 4 February 2021

DOI: 10.1111/aogs.14125

AOGS Acta Obstetricia et Gynecologica Scandinavica

ORIGINAL RESEARCH ARTICLE

Evaluation of repeat testing of a non-sequencing based NIPT test on a Finnish general-risk population

Filip Karlsson[1] | **Tarja Ahola**[1] | **Jenny Dahlberg**[1] | **Lawrence Prensky**[2] | **Helena Moilanen**[3] | **Heli Spalding**[4]

[1]PerkinElmer Inc., Vanadis Diagnostics, Stockholm, Sweden

[2]PerkinElmer Inc., Waltham, MA, USA

[3]Nordlab, Oulu, Finland

[4]Obstetrics and Gynecology, Terveystalo Healthcare Services, Helsinki, Finland

Correspondence
Filip Karlsson, PerkinElmer Inc., Vanadis Diagnostics, Vetenskapsvägen 10, Stockholm, Sweden.
Email: filip.karlsson@perkinelmer.com

Abstract

Introduction: To evaluate the effect of repeating test failures using an automated, non-sequencing based non-invasive prenatal testing test on a general-risk population in Finland.

Material and Methods: A total of 545 samples from women who represent the average-risk population in Oulu, Finland were analyzed with Vanadis® non-invasive prenatal testing. Repeat testing of test failures was performed using a second sample. Results before and after repeat testing were compared with the reference outcome, as determined by clinical examination of neonates.

Results: There were eight test failures after first-pass analysis, representing 1.5% of samples (95% CI 0.6%-2.9%). Seven out of eight failures could be resolved by analysis of a second sample, thereby reducing the test failure rate from 1.5% to 0.2% (95% CI 0.0%-1.0%).

Conclusions: Repeating test failures with a second plasma sample could significantly reduce the effective failure rate, thereby providing a way to effectively minimize test failures and further improving clinical utility and test performance.

KEYWORDS
cell-free DNA, non-invasive prenatal screening, non-invasive prenatal testing, prenatal screening, prenatal testing, test failure, Vanadis

1 | INTRODUCTION

Since the introduction of cell-free DNA (cfDNA) based non-invasive prenatal testing (NIPT) over the last decade, numerous studies have reported on its screening performance in terms of detection rate and false-positive rate.[1]

Another NIPT performance metric that is sometimes overlooked is the test failure (or no-call) rate. In addition to being confusing for the woman and physician receiving such a result, failures also negatively affect the actual sensitivity or detection rate of the test because a proportion of failures can be assumed to be trisomy cases. This can be illustrated by considering a hypothetical NIPT test with 100% Trisomy 21 (T21) detection rate at varying failure rates, similar to that previously presented by Yaron.[2] For a failure rate of 1%, 5%, or 10%, the corresponding actual detection rate of the screened population will deteriorate to 99%, 95%, and 90%, respectively, if

Abbreviations: cfDNA, cell-free DNA; NIPT, non-invasive prenatal testing; T13, trisomy 13; T18, trisomy 18; T21, trisomy 21.

assuming the same incidence of T21 among failures as for reported results. Hence, the test failure rate will influence the actual screening performance negatively and should be kept as low as possible.

As a real-world example of the same point, Norton et al reported 100% detection rate for T21 (38 of 38) from 15 841 women with a successful cfDNA-based NIPT result.[3] There were three additional T21 cases among the 488 (3%) women who received a failed NIPT result. Therefore, the detection rate of the population would be decreased from 100% to 93% when considering the failed cases as well.

In a 2016 review of NIPT studies reporting test failures, Yaron analyzed 20 studies totaling 429 624 assayed patients.[2] Failure rates were shown to differ with the NIPT technology used, with 1.6% failures for massive parallel sequencing methods, 3.6% for chromosome-specific sequencing methods, and 6.4% for single nucleotide polymorphism-based methods.

Similarly, a 2017 NIPT meta-analysis by Gil et al showed a high variability of reported failure rates, ranging from 0.1% up to 8.8%, although the authors were not able to draw conclusions on differences in failure rate based on method used.[1] A 2018 systematic review of NIPT failure rates from 30 studies, stratified by Western and Asian studies with initial testing and Western studies after repeat testing, revealed failure rates of 3.3%, 0.6%, and 1.2%, respectively.[4]

One potential way of reducing the failure rate is to repeat test failures by analyzing a second sample. The second sample can either be collected in conjunction with the first sample or from a redraw upon receiving the failed result. Directly collecting two tubes minimally decreases the turnaround time because the sample is already available in the laboratory but comes at the cost of having to collect two samples from all patients. Redrawing a second sample after receiving a failed result avoids this, but it will increase the turnaround time and one study reported that only 56% of patients show up for a redraw.[5]

The ability to resolve a test failure with a second sample likely depends on the source of the failure. The possibility to resolve failures due to inherent sample characteristics such as low fetal fraction or abnormal total cfDNA amounts is likely lower than for other types of assay failures not directly related to sample characteristics, such as technical failures affecting data quality.

The Vanadis NIPT test used in this study has previously been validated with a low failure rate on a high-risk population.[6] The present study was carried out to evaluate test performance in a general-risk population. In addition, a screening strategy to reduce the test failure rate through repeat testing was evaluated.

Key message

Test failures in non-invasive prenatal testing have the potential to cause clinical concerns. Our study shows that performing a repeat test on a second sample can significantly reduce test failures.

2 | MATERIAL AND METHODS

2.1 | Study population

A total of 545 samples collected between May 2016 and September 2018 at maternity clinics in Oulu, Finland were assayed. Samples were collected in association with routine maternal blood testing, including, but not limited to, first-trimester screening. Two blood samples (10 mL) were drawn from women who consented to donate extra samples for the study. These women were between the ages of 18 and 50 years; and were between 9 and 40 weeks pregnant. Women were not excluded for any other reason. Participants were not selected based on prior risk.

2.2 | Sample collection

Samples were collected into Cell-Free™ DNA BCT tubes (Streck) and shipped by courier to the Vanadis Diagnostics Laboratory. Upon arrival, blood samples were processed to plasma using a double centrifugation protocol of 1300 *g* for 30 minutes followed by 2400 *g* for 20 minutes. The plasma fraction was transferred to an intermediate container following the first centrifugation step and transferred to a storage tube after the second centrifugation step. The plasma fraction was extracted within 5 days of blood draw and stored at −80°C until processing at the Vanadis Diagnostics test facility. Processing and analysis of plasma samples using Vanadis NIPT (PerkinElmer) was performed blinded to the birth outcomes.

2.3 | Test method

The Vanadis NIPT assay relies on a series of enzymatic steps that specifically generate labeled rolling circle replication products from chromosomal cfDNA targets, as previously described.[7] Automated extraction of cfDNA from plasma was performed using the Vanadis Extract® platform, followed by continued processing on the Vanadis Core® platform to generate labeled rolling circle replication products that were imaged and counted using a Vanadis View® instrument.

Automated data analysis and quality assessment were performed, and chromosomal ratios were calculated for all approved samples as described previously.[7] The quality assessment process approves or disqualifies samples based on several metrics related to either data quality or the total number of counted objects, which is related to the cfDNA concentration of the sample (not to be mistaken for the fetal fraction of the sample).

The approved samples were classified into low or high risk using a *z* score approach based on each normalized chromosomal ratio and the sample-specific standard deviation. Trisomy 13 (T13), Trisomy 18 (T18), and T21 classification was performed using default *z* score cut-offs of ≥3.50 for T21 and ≥3.15 for T13 and T18.

Test failures were repeated using a second sample. If the failure mode of the first sample was due to high cfDNA concentration of the sample, then a 10-fold dilution step of the second plasma sample was performed before repeat testing. The samples were diluted into a 1× phosphate-buffered saline solution (Sigma-Aldrich, #806552).

Vanadis NIPT results were compared with neonatal examination results for the 507 samples in order to evaluate false-positive rates. All 534 samples were included for evaluation of failure rates before and after repeat testing. All confidence intervals were calculated using the Clopper-Pearson method.

2.4 | Ethical approval

Protocols used for sample collection were approved by the Northern Ostrobothnia Hospital District regional ethics review committee (EETTMK 123/2015) on 18 January 2016 and informed consent was obtained from all participants. The study was performed in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration.

3 | RESULTS

Of the 545 samples analyzed, 11 were excluded from analysis because of incorrect barcoding of the sample. Among the remaining 534 samples, there were seven samples with first-trimester screening risk ≥1/250. One of these seven had Cornelia de Lange syndrome. The remaining six were normal on follow up, confirmed either by cytogenetic testing or examination after birth. There were also six twin pregnancies and seven in vitro fertilization pregnancies.

In all, 507 samples were followed to birth and concordance was determined by clinical examination of the neonates. All 507 neonates were considered healthy upon birth. Among the 27 cases that could not be followed to birth, 23 had a pregnancy loss or an abortion due to a fetal anomaly. For the remaining four, birth outcome data could not be obtained for other reasons.

The median maternal age of the 507 samples followed to birth was 29 years (interquartile range [IQR] 25-33). All samples were collected between 9^{+0} and 16^{+1} weeks of gestation. The median gestational age was 10^{+5} weeks (IQR 10^{+2}-11^{+2} weeks) and the median maternal weight was 67 kg (IQR 59-76 kg). Patients were not supplied with test results from this exercise.

3.1 | Test performance

In the group of 507 samples with outcome information available, two were classified as high risk (0 for T21, 2 for T18, and 0 for T13). The *z* scores for these two T18 high-risk samples were 3.3 and 4.8, respectively. The overall false-positive rate of this group after first-pass analysis was 0.4% (95% CI 0.1%-1.6%).

Among the group of 27 samples without outcome information, one was classified as high risk for T21 with Vanadis NIPT, with pregnancy loss being the reason that no birth follow-up data were available (*z* score of 4.9).

Among the 534 samples assayed, there were eight assay failures after first pass analysis resulting in a first failure rate of 1.5% (95% CI 0.6%-2.9%). Six samples failed quality assessment metrics associated with counts being out of bounds. Of these, five were above the upper count limit, likely due to abnormally high cfDNA amounts, and one failed because of insufficient counts, likely caused by low cfDNA amount in the sample. The two remaining failures could not be related to cfDNA amount; one was caused by incoherent chromosomal ratios and the other by a high measurement variation between individual images of the well. Table 1 shows the demographics of this cohort.

The eight test failures were re-processed using a second plasma sample, five of which were diluted before re-processing based on their failure mode being related to abnormally high cfDNA concentration.

A result was obtained for seven of eight (87.5%) re-processed test failures and all of these results were accurate. The remaining failure

TABLE 1 Demographics of patients with test failure

	Maternal age	BMI (kg/m^2)	Weight (kg)	Gestational age (wk)[a]	Pregnancy outcome	FTS result T21 1/NNN	FTS result T18 1/NNN	Reason for test failure
Patient 1	23	23.6	59	10^{+3}	Miscarriage at 13 weeks	N/A	N/A	Counts above upper limit
Patient 2	28	19.7	55	11^{+4}	Healthy boy	5800	N/A	Image measurement variation above limit
Patient 3	27	27.8	72	10^{+4}	Healthy girl	100 000	100 000	Counts above upper limit
Patient 4	25	23.8	68	10^{+0}	Healthy boy	45 104	100 000	Counts above upper limit
Patient 5	35	22.1	64	11^{+1}	Healthy girl	21 144	100 000	Counts above upper limit
Patient 6	33	20.9	57	10^{+1}	Healthy boy	320	6731	Insufficient counts
Patient 7	21	31.1	93	11^{+1}	Healthy boy	20 193	100 000	Incoherent ratios
Patient 8	27	29.7	79	10^{+4}	Healthy boy	100 000	100 000	Counts above upper limit

Abbreviations: BMI, body mass index; FTS, first-trimester screening.

[a]Based on ultrasound dating.

was caused by insufficient counts for analysis, which was the same failure mode as for the original sample. Hence, the number of failures was reduced from eight to one, resulting in an effective failure rate of 0.2% (95% CI 0.0%-1.0%) in the tested population of 534 samples.

4 | DISCUSSION

In this study we found that test failures could be reduced by repeat testing. Test failures were reduced from eight to one by analyzing a second sample.

The majority (five out of eight) of the failures in this study were related to high cfDNA concentration. A high success rate for resolving these could be achieved by diluting the sample before repeat testing, as it can be assumed that the second sample has the same cfDNA concentration as the first, given the fact that both samples were collected at the same time.

The only failure after repeat testing was caused by a low cfDNA concentration, which was the same failure mode as the original sample. Low DNA concentration failures are more difficult to resolve because the second sample will likely also have a low concentration. As there were no apparent differences between the demographics of these samples and the rest of the cohort, we cannot explain why some samples have high or low cfDNA concentrations, nor can we explain why this one sample failed twice. However, we are hopeful that other studies with larger populations will be able to answer these questions.

Both the magnitude and types of failures can be heavily dependent on the test methodology used, so it is not clear if the findings from this study can be directly applied to other test methodologies. Even so, the findings herein indicate that the failure rate for the Vanadis NIPT test used in this study can be kept minimally low using a repeat-testing approach, even though a larger sample size is needed for more accurate quantification of the failure rate before and after repeat testing.

It should be noted that the benefits of reduced failure rate through repeat testing does require that either two blood tubes be collected from each woman to be tested or that a re-draw is requested after analysis. Another potential option is to recommend invasive follow-up testing directly on all test failures but doing this would effectively decrease the positive predictive value of the test, especially if the failure rate were high.

The first pass failure rate of 1.5% in this study was higher than what has previously been reported for this technology,[6,7] but still significantly lower than most other technologies and is comparable to that of massive parallel sequencing-based NIPT methods, which showed the lowest average failure rate based on a meta-analysis of NIPT failure rates.[2] By also employing a repeat-testing approach with this technology, the effective failure rate in this study was reduced by 87% to 0.2%. Consequently, the clinical utility and sensitivity of the test were improved, and difficult counseling decisions associated with handling and reporting test failures could be avoided. Another benefit of a low failure rate can be less invasive testing needed, as one approach of resolving failed samples is to follow up with cytogenetic testing from invasive sampling, so lessening pregnancy losses, parental stress, financial costs, and clinical burden.

5 | CONCLUSION

Repeating test failures with a second plasma sample could significantly reduce the effective failure rate, thereby providing a way to effectively minimize test failures and further improving clinical utility and test performance.

ACKNOWLEDGMENTS

We wish to thank the midwives in the city of Oulu, Finland who recruited the women that participated in the study.

CONFLICT OF INTEREST

FK, TA, JD, and LP are employed by PerkinElmer.

ORCID

Filip Karlsson https://orcid.org/0000-0002-0860-1500
Lawrence Prensky https://orcid.org/0000-0003-4427-9122

REFERENCES

1. 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;50:302–314.
2. Yaron Y. The implications of non-invasive prenatal testing failures: a review of an under-discussed phenomenon. *Prenat Diagn*. 2016;36:391–396.
3. Norton ME, Jacobsson B, Swamy GK, et al. Cell-free DNA analysis for noninvasive examination of trisomy. *N Engl J Med*. 2015;372:1589–1597.
4. Palomaki GE, Kloza EM. Prenatal cell-free DNA screening test failures: a systematic review of failure rates, risks of Down syndrome, and impact of repeat testing. *Genet Med*. 2018;20:1312–1323.
5. Dar Pe'er, Curnow KJ, Gross SJ, et al. Clinical experience and follow-up with large scale single-nucleotide polymorphism-based noninvasive prenatal aneuploidy testing. *Am J Obstet Gynecol*. 2014;211:e1–e17.
6. Ericsson O, Ahola T, Dahl F, et al. Clinical validation of a novel automated cell-free DNA screening assay for trisomies 21, 13, and 18 in maternal plasma. *Prenat Diagn*. 2019;39:1011–1015.
7. Dahl F, Ericsson O, Karlberg O, et al. Imaging single DNA molecules for high precision NIPT. *Sci Rep*. 2018;8:4549.

How to cite this article: Karlsson F, Ahola T, Dahlberg J, Prensky L, Moilanen H, Spalding H. Evaluation of repeat testing of a non-sequencing based NIPT test on a Finnish general-risk population. *Acta Obstet Gynecol Scand*. 2021;100:1497–1500. https://doi.org/10.1111/aogs.14125