Introduction

We are thrilled to share with you our new whitepapers, highlighting the much-debated topic, “fetal fraction”. Fetal fraction is the proportion of DNA in the maternal blood that came from the pregnancy and has been used by some laboratories to determine if they can have confidence in their results. However, there are significant factors to consider when determining its merit, including assay precision, the accuracy of a reported fetal fraction metric and the implications of using a fetal fraction cutoff.

In the enclosed publication, we present the science, the data and the impact of using the fetal fraction metric within a clinical context. We hope you’ll find this series informative and of great value to your laboratory, your practice or your pregnancy.

<table>
<thead>
<tr>
<th>TABLE OF CONTENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Understanding Fetal Fraction in cfDNA Screening .......................................................... 3</td>
</tr>
<tr>
<td>Why Fetal Fraction is Considered Important ................................................................. 6</td>
</tr>
<tr>
<td>The importance of Precision and Fetal Fraction in NIPT ................................................ 8</td>
</tr>
<tr>
<td>Understanding the Implications of No-Calls ................................................................. 11</td>
</tr>
</tbody>
</table>
Fetal fraction (FF) is the percentage of all DNA in maternal blood that has come from the pregnancy. It has often been considered an important quality metric on a single sample level because when the FF of a sample is low, laboratories often have less confidence in determining a sample’s affected status. (To learn more, read our white paper entitled “Why Fetal Fraction is considered important”). As a result, many laboratories use a FF cut-off (often at 4% FF) where they do not report out results for samples below this percentage due to their lack of confidence in their test’s ability to provide an accurate answer.

Although this may appear to make sense in theory, there are a number of issues with using a FF cut-off that need to be considered:

1. The reported FF for all companies is only an estimate
2. How to view fetal fraction in a proper context
3. The implications to patients by discarding samples due to a low FF is significant

1. REPORTED FETAL FRACTION DOES NOT CORRECTLY REFLECT THE REAL FETAL FRACTION

The reported FF that is provided on a given sample is a measured value. Unfortunately, this measured value is only an estimate of the actual (true) FF. Although methods for measuring FF can be good at estimating the average true FF of a large group of samples, they are notoriously poor at estimating true FF for individual samples.

For example, the College of American Pathology published a paper in 2020 that looked at many aspects of a laboratory’s NIPT capabilities.¹ One of the factors they examined was the reported FF that was provided for a sample.

When comparing the results from the 56 laboratories that took part in this aspect of their study, they found that for an individual sample, laboratories gave a reported FF that could range over 18% (i.e. for a given sample “X”, one lab reported the FF as 7% and another reported the FF as 25%, while the other labs reported FFs anywhere in between). This study showed that the reported FF for samples are, at best, only estimates with likely significant deviation from the true FF.
Another example of how imprecise a reported FF can be, is seen in the letter by Takoudes et al.² Here, they showed how a non-pregnant woman (so a true FF of 0%) was given a reported FF of 4.3% and another non-pregnant woman was given a reported FF of 3.9%.

So why are these reported FF often incorrect? If you look at the different methodologies that are used to try to determine the FF of a sample, the literature shows that these methods report a FF that is typically within +/- 3% of the true FF.⁴ This means that when you see a reported FF value of 5% it means that it originated from a sample with a true FF in the range of 2 - 8% (with a 95% probability), as illustrated in Figure 1. Therefore, it is not surprising that the reported FF is often incorrect.

Additionally, in an article from 2015, Wright et al.³ found that 2.99% of all samples have a reported FF below 4%, while the percentage of the samples which would actually have a true FF below 4%, was only 0.37%. This means that at least 88% of the samples with a reported FF below 4% are discarded unnecessarily and have a true FF above 4% and could thus be classified appropriately by the test.

This shows that having a low true FF is rare, whereas having a low reported FF is common. This imprecision in FF measurement means that the majority of samples discarded due to a low reported FF will have been discarded unnecessarily. The consequence of this is a significant increase in the “no-call” rate.

2. HOW TO VIEW FETAL FRACTION IN A PROPER CONTEXT

Although understanding the difference between true FF and reported FF is important, there is another aspect surrounding FF that is important to understand as well. When trying to determine whether the true FF of a sample is at a level that a test feels confident in providing a result, you must also look at the other side of the equation and see whether the test itself has the ability to provide a result at a given FF. In other words, there are some NIPT tests that can confidently provide results at lower true FFs than other tests. This is one of the reasons why different technologies have different FF thresholds: Some tests can differentiate between normal and abnormal samples at lower FF, while some can only do so at higher FF.

The way these tests are able to do this, is by using a technological approach that provides a greater precision than another NIPT test. Tests that have high precision are less sensitive to the FF because they can confidently resolve samples at most FFs.⁴ Whereas tests with lower precisions rely heavily on the reported FF cutoff in order to avoid making decisions about samples for which they are not confident of calling correctly.

(To learn more, read our white paper entitled “The importance of Precision and Fetal Fraction in NIPT”.

3. IMPLICATIONS OF USING A REPORTED FETAL FRACTION CUT-OFF

The final issue revolves around how using a reported FF cut-off will cause an increase in the no-call rate. There are three major implications of this:

- The effect on the claims of the laboratory regarding the test performance.
- Not using a reported FF cut-off value to “improve performance” actually provides a benefit to the effective performance of the test.
- Having a high no-call rate means that many women who perform the test would be denied an NIPT result due to a low reported FF. But by removing the FF cut-off, the vast majority of these women could then receive results and consequently avoided the anxiety and stress associated with making decisions about samples for which they are not confident of calling correctly.

(To learn more, read our white paper entitled “Understanding the implications of “no-calls””.)
References


Authors

Lawrence Prensky
PerkinElmer Inc.,
Waltham, MA

Fredrik Persson
Vanadis Diagnostics
PerkinElmer Inc.,
Sollentuna, Sweden
We wanted to briefly explain why fetal fraction (FF) tends to be a focus for some laboratories. In our other papers, we will explain why this emphasis on a reported FF is excessive.

Fetal fraction (FF) is often considered a limiting factor in the clinical accuracy and usability of non-invasive prenatal testing (NIPT). Due to this belief, many laboratories do not provide results in samples that they think have a low FF, since they feel that their test may be inaccurate with these samples. The reason for their concern is that when the amount of FF is low, it is difficult for them to distinguish between samples from a pregnancy with, say, Down syndrome versus samples from pregnancies without Down syndrome. This can be illustrated by looking at the following example:

Let’s pretend we’re looking at how tall girls are. If we collect data on all the 6-year-old girls out there, we’d find that their heights vary, but tend to fall around an average of 45 inches. Some 6-year-old girls are taller and some are shorter, but most are around the average. When we plot the number of girls at each height on a graph, it forms a shape called a “Bell Curve”. Now, let’s look at the heights of 16-year-old girls. We will see the same curve, but the average height will be greater (64 inches), as shown in Figure 1. So if I told you that there was a girl who was 49 inches tall, and you had to decide whether this was a 6 year old or a 16 year old, you could be fairly confident in saying she is a 6 year old.

But now let’s compare 6-year-old girls with 10-year-old girls (average of 54 inches). It would be harder to say if a girl who is 49 inches tall is a tall 6-year-old or a shorter 10-year-old because there is more overlap between girls at this height (see Figure 2). In this case, you would just say “I don’t know” and avoid guessing. It is this overlapping of curves, and the fear of giving an incorrect result, that is similar to what is happening with FF and NIPT testing.

---

*Figure 1. The height distributions for 6 and 16 year old girls (schematic illustration).*
However, in figure 4, you can see the expected curves when a low FF is chosen (3%). In the highlighted region, there is an overlap between the affected curve and the unaffected curve. Which means that if a result landed under the “affected” curve, we’d expect the pregnancy to be unaffected. If it landed under the “unaffected” curve, then we’d expect the pregnancy to be affected. This is because there is a very clear differentiation between the two curves.

This is why FF has often been considered so important by laboratories. If they didn’t exclude those samples, then their test would be wrong more often and their sensitivity, specificity, positive and negative predictive values would suffer. So it is understandable why a laboratory might believe it is necessary to use a FF cut-off and why they believe that knowing the FF of a given sample would be useful. But all is not as it seems. Additional information is required (such as an understanding of how FF is measured; the difference between a reported fetal fraction and the true fetal fraction; and the imprecision of the NIPT assay being employed) to fully understand why the current emphasis on the reported FF is overstated.

**Authors**

Lawrence Prensky  
PerkinElmer Inc.,  
Waltham, MA, USA

Fredrik Persson  
PerkinElmer Inc.,  
Sollentuna, Sweden
The way an NIPT assay is able to determine if a sample came from an affected pregnancy or not, is by looking at the chromosomal ratio. The chromosomal ratio is a measurement of how much DNA from a specific chromosome there is compared to the amount of DNA from a reference chromosome. For an unaffected pregnancy this ratio is expected to be around 1. For an affected pregnancy the ratio is expected to be around 1+FF/2, where “FF” stands for Fetal Fraction and represents the proportion of DNA in mom’s blood that came from the pregnancy (typically around 10%). In pregnancies with a trisomy (three copies of one chromosome instead of the typical two copies), the DNA that originates from the pregnancy should contain 50% more material from the trisomic chromosome (such as chromosome 21 in the case of Down syndrome) as compared to the reference chromosome, yielding a ratio around 1+FF/2.

Background

Like measuring most things in nature, the measurement of DNA is imperfect, so not all samples will get a perfect chromosomal ratio of 1 or 1+FF/2. Some would get slightly lower or slightly higher ratios, but all will still be centered around the expected value. How much the samples vary around the expected value is described by the standard deviation (SD), which depends on the precision of the measurement. If we plot the number of samples at each chromosomal ratio on a graph, it forms a shape known as a “Bell Curve” or “Gaussian distribution”. The width of this curve is set by the SD and is determined by the precision. Figure 1 (top left) illustrates the Bell curves for unaffected pregnancies and affected pregnancies with a FF of 10%. In addition to visualizing the samples as chromosomal ratios, they are also shown here in terms of what is called “Z-scores”. The Z-score can be thought of as how far away from the average of unaffected pregnancies, a specific sample is located. So, the farther from the average a point is located, the greater the Z-score. And the greater the Z-score, the less likely it is from an unaffected sample. Z-scores are measured in units of “number of SDs”.

The advantage of using Z-scores is that they already compensate for different variability/precision in different assays, so regardless of the test, 99.9% of all unaffected pregnancies should have a Z-score less than three. This is why three is commonly used as a cut-off for separating between calling low / high chance of having an affected pregnancy. In Figure 1 (top), we can clearly see that there is a relatively large separation between the unaffected and affected pregnancies when the FF is 10%, making it easy to distinguish between an unaffected and affected pregnancy at this FF.
The Concern

The challenge arises when the FF becomes lower, meaning that the affected curve moves closer to the unaffected curve and can start to overlap, as shown in Figure 1 (bottom). Here, at a FF of 4%, it is clear that the curves can start to overlap and some of the affected pregnancies can not be detected using a Z-score of three. With this increased overlap, the confidence a laboratory has in reporting out a result decreases since they know that they can be giving incorrect results. In order to try to avoid giving incorrect results, many laboratories choose not to report out results when the FF is low. This is why laboratories are often quoting a measured FF on their reports and choose a "FF cutoff", where they withhold a result if the measured FF is below a certain FF%.

The benefit of this to the laboratory, is that it allows the laboratory to only report out results that they have a higher degree of confidence in. The problem with this, is that the "estimated" FF for a given sample that is measured is often significantly different from the true FF, leading to the vast majority (~90%) of women with a low estimated FF not getting a result when they actually had a sufficiently high true FF to get a reliable result. Considering that it is a relatively large proportion of samples (typically 1-5%) that are discarded as "no-calls" due to this cut-off in reported FF, it is striking that the vast majority of these samples are basically collateral damage due to the challenges associated with accurately measuring the FF for individual samples. Additionally, a significant fraction of the few samples that do have a low FF will have a measured FF that looks normal, meaning that the added confidence of a normal FF is in part imaginary.
Precision to the Rescue

So what are we to do? We have just shown that knowing the true FF is important for NIPT, but we recognize that there are limitations that prevent any of us from determining what the true FF for a given sample actually is. The solution is in the width of the chromosomal ratio distribution, i.e. the precision of the test and thus the SD. Remember that SD represents how big the spread of values is around the average (the width of the curve). So the closer most of the samples are to the average, the smaller the SD. The statistical term to describe how close samples are to each other is called precision. The more precise a test is, the closer the samples are to the average, so the smaller the SD would become.

Figure 2 (top) again shows the situation where a significant fraction of the affected samples cannot be called correctly with the Z-score cut-off of 3. However, for a test with a higher precision (meaning a lower SD) the chromosomal ratio curves becomes narrower (Figure 2 bottom left). Since Z-scores are measured in units of SDs, this leads to the Z-scores for the affected samples being shifted to higher values, leading to less overlap between the affected and unaffected z-score curves (as shown in Figure 2 (bottom right)). This means that although the true FF is still an important component for NIPT performance, its importance can be minimized by using a test with greater precision, with the added benefit of being able to report results to all women.

So how do we increase a test’s precision? There are basically three key things that influence the precision:

- **Statistics** - Increase the number of DNA fragments that are counted from the relevant chromosomes. The more pieces you count, the more precise the chromosomal ratio will become. For NIPT this means that targeted methods have an advantage over whole genome methods since all of the counts in the targeted method will be relevant, while many counts in the whole genome method will often be irrelevant (since many are from chromosomes that are not being reported on or are not from regions unique to a given chromosome).

- **Amplification** - If the quantity that is measured is amplified, the variability is also amplified. For example, most NIPT tests use PCR (polymerase chain reactions) to make millions of copies of each piece of cfDNA they find in the blood. However, PCR can not faithfully copy each piece the exact same number of times (i.e. there is variation during the amplification). Therefore, tests that use PCR tend to be less precise because PCR can increase variability by 30-50%.5

- **Automation** - Every sample and the steps it goes through in the workflow should be as reproducible as possible to ensure that the variability is minimized. By increasing the degree of automation and thereby lowering the degree of manual processes, the precision also increases.

As a consequence, the focus should be on developing a more precise test to mitigate the FF component and still provide accurate results, instead of relying on cut-offs for a reported FF that is inherently inaccurate. The added clinical benefit to this is that the test does not need to use a FF cut-off and deny many women accurate test results. These women could then avoid follow-up invasive procedures, thus lessening pregnancy losses, parental stress, financial costs and clinical burden.

References


Authors

Lawrence Prensky
PerkinElmer Inc.,
Waltham, MA, USA

Fredrik Persson
PerkinElmer Inc.,
Sollentuna, Sweden

For a complete listing of our global offices, visit www.perkinelmer.com/ContactUs

Copyright ©2021, PerkinElmer, Inc. All rights reserved. PerkinElmer® is a registered trademark of PerkinElmer, Inc. All other trademarks are the property of their respective owners.
Any time a woman undergoes a non-invasive prenatal screening test, there is a chance that she finds herself in an adverse situation based upon her results.

For the sake of our discussion, let’s focus on the adverse situations where a woman is given a result that is classified as false negative, false positive or as a no-call. The adverse outcomes for false negatives (the unexpected birth of a child with aneuploidy) and false positives (unnecessary invasive procedures and the additional risk of miscarriage) are significantly lower with NIPS than other prenatal screening tests. However, the risk for receiving a no-call result is significantly higher with NIPS. **No-call results are when the laboratory does not or can not provide a result on a given sample.**

**Potential Reasons for a No-call**

- **Technical reasons:** such as the tubes arrive broken, there are user errors, the machine stops working, or for those methods using sequencing - poor sequencing quality.
- **Sample reasons:** such as having too much or too little cell-free DNA (cfDNA) in the sample.
- **Biological anomalies:** such as an uncommon chromosomal error in the fetus or placenta; The mother could have a chromosomal anomaly herself; or even possibly cancer.
- **Low fetal fraction (FF):** where the results come back with a reported fetal fraction below the lab’s cutoff. In this scenario, the reported amount of fetal DNA in the maternal blood is too low for the lab to have confidence in reporting out a result.

**Consequences of Receiving a No-call Result**

So what happens to these women when they receive a no-call result? In most cases, it depends upon the reason for the no-call. All samples can get offered redraws to try to resolve the no-call result. Although some causes of no-calls are likely to be resolved at a high rate with redraws (e.g. around 90% for technical issues), not all no-calls can be settled in this manner. There is also the inconvenience of going back to the lab to get redrawn (plus having to contend with a needle again). That may not seem like much of an adverse situation, but if a woman is unable/unwilling to undergo a redraw, or the redraw did not resolve the no-call, then there are two potential consequences: no results or unnecessarily delayed results. In which case, maybe something gets missed or they don’t get the opportunity to make informed decisions about the pregnancy.

When it is something other than a technical or sample reason, it is less likely to be resolved and thus women are often offered diagnostic testing (such as a CVS or amniocentesis). There are multiple guidelines from organizations such as SMFM and ACOG, that recommend that diagnostic testing be offered to women who receive no-call results.
If a woman is offered a diagnostic test, then there could be a number of adverse situations:

- You would likely be increasing the anxiety and fear in these couples. Getting unexpected results, especially ones that could be associated with pregnancy or maternal concerns, will scare people. If you then add on the anxiety that can be associated with performing a diagnostic test, you can see why a no-call is not an optimal result for couples.
- There is also a financial burden. Although most insurances, private or state, will cover the costs of follow up for this woman, that is not always the case. There could therefore be an unnecessary financial cost to the woman or at the very least, to the medical system.
- When a no-call result is received by the medical provider, that provider then has a lot of work to do. They need to understand why there was a no-call, contact the woman, explain the lack of results to the woman, discuss next steps (including diagnostic testing), deal with the possible anxiety and fear from that woman, arrange or perform the diagnostic test, and then handle those results. It is a lot of work and the clinical burden can be significant.
- Finally, if these women do undergo a diagnostic test, then there is that associated risk of miscarriage and losing what is likely a healthy baby.

**How Do We Minimize These Risks?**

- Since ~90% of no-calls get resolved when retested due to a technical issue, we can just redraw these women. However, studies have shown that 40-50% of women called for a re-draw, will not actually be redrawn. This risk can be mitigated by drawing two tubes initially so an extra tube is on hand if needed.
- For sample issues, if there is too much or too little cfDNA, sometimes a lab can dilute or concentrate the sample, but this can sometimes require a second tube of blood as well.
- It would be best not to minimize biological issues, since these are cases that often benefit from additional investigation.
- For FF cutoff, we can just remove the cutoff altogether and provide all these women with results. While some people state that FF is a critical metric of the sample, we would suggest reading our other papers to learn why 1) the FF is often incorrect; 2) the chance of getting a true FF below the cutoff is actually very small, and; 3) removing a FF cutoff actually provides a benefit to the effective performance of the test.

So what are the benefits of removing a FF cutoff? To begin with, you get fewer no-calls, which means that you minimize undue stress and anxiety on parents, along with causing less clinical burden, less financial burden and fewer invasive procedures, which means fewer pregnancy losses.

**Implications in the Real-world**

In the United Kingdom, historically, there are around 700,000 babies born every year. If we say that the incidence of Down syndrome is approximately 1/700, then there would be 1,000 women carrying a Down syndrome pregnancy in the United Kingdom every year.

Let's now compare two hypothetical NIPT tests for this population. The first is what many labs currently use, with a 4% FF cutoff and average precision – meaning that the test can resolve samples at 4%. The second is the same test but with the FF cutoff removed.

If we extrapolate the data from Wright et al, you can see that in those 1,000 women who are carrying a Down syndrome pregnancy, the current test would give correct results 966 times. 30 of these samples would have a reported FF below the 4% cutoff and would be classified as a no-call. And four samples would get False Negative (FN) results.

<table>
<thead>
<tr>
<th></th>
<th>Test #1</th>
<th>Test #2</th>
</tr>
</thead>
<tbody>
<tr>
<td>T21 Samples</td>
<td>966</td>
<td>995</td>
</tr>
<tr>
<td>No Call Due to FF Cutoff</td>
<td>30</td>
<td>0</td>
</tr>
<tr>
<td>True Positive</td>
<td>966</td>
<td>995</td>
</tr>
<tr>
<td>False Negative</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Unexpected T21 Births</td>
<td>17-24</td>
<td>5</td>
</tr>
<tr>
<td>All Samples</td>
<td>No Call</td>
<td>21,000</td>
</tr>
</tbody>
</table>

If we remove the FF cutoff, then all 30 of these women would get results. So there wouldn’t be any no-calls due to the cutoff (there would likely still be no-calls for other reasons, but those wouldn’t be any different between the two tests). Therefore, by not having any no-calls, this would mean that 995 out of 1000 women would be given correct results and five samples would be given FN results. Therefore, 29 out of the 30 women who didn’t get results, would have received correct results.

Someone could say that getting those additional 29 correct results at the expense of 1 FN is not worth it. That even having that 1 FN in the cohort of 700,000 women is not acceptable. But what this needs to be balanced against is that even though the recommendations are for no-call women to get a diagnostic test (so that these 30 samples would have been caught anyway), about 45-68% of women do not go on to have a diagnostic test. So of these 30 women who are carrying a Down syndrome fetus, at least 13-20 will skip an amnio and ultimately give birth to a child with Down syndrome. Meaning that for test #1 – when we factor the 4 FN in, there will be 17-24 unexpected Down syndrome births, while for test #2 – there would only be five.

Those are some of the outcomes that could occur to the 1000 samples that are from women who are carrying a Down syndrome pregnancy. Additionally, there are ramifications to the 699,000 samples that are from pregnancies without Down syndrome. In that group, there would be almost 21,000 women who would be receiving results, since their samples wouldn’t be discarded due to having a reported FF below the cutoff.
The final implication would be that by removing the FF cutoff, we ensure that an additional 21,000 women are provided a highly reliable NIPT result each year in the United Kingdom. By avoiding a no-call designation in these 21,000 women, we would be sparing them (and their providers) from the stress, cost and clinical burden associated with invasive procedures. On top of that, depending upon which rates you want to use, you would also be preventing the 21-210 miscarriages7-9 that would have occurred by doing that procedure.

And that is just for the United Kingdom. If you looked at Europe’s 4 million babies a year,9 you are talking about 120,000 women with no-calls because of this cut-off, which could lead to 120-1,200 miscarriages every year due to an unnecessary invasive procedure.7-9

Minimizing adverse situations is better for everyone and can be accomplished simply by decreasing the number of no-calls a laboratory reports out. Doing so would help everyone by reducing redraws, anxiety, costs, clinical burden and most importantly, pregnancy loss. It can also reduce the chance of having an affected child born unexpectedly. We are able to minimize these adverse outcomes just by removing the FF cutoff and allowing a precise test to give accurate results to more women.

References


Authors

Lawrence Prensky
PerkinElmer Inc., Waltham, MA

Fredrik Persson
PerkinElmer Inc., Sollentuna, Sweden